AFRL-RH-WP-TR-2012-0114

ANALYST PERFORMANCE MEASURES
VOLUME III: INFORMATION QUALITY TOOLS FOR PERSISTENT SURVEILLANCE DATA SETS

John Talburt
Serhan Dagtas
Mariofanna Milanova
Mihail Tudoreanu
University of Arkansas at Little Rock
2801 South University Avenue
Little Rock, AR 72204

Brian Tsou
Forecasting Division
Human-Analyst Augmentation Branch

November 2011
Final Report

Distribution A: Approved for public release; distribution is unlimited.

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING
HUMAN EFFECTIVENESS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2012-0114 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signed//

BRIAN TSOU
Work Unit Manager
Human-Analyst Augmentation Branch

//signed//

LOUISE A. CARTER, Ph.D.
Chief, Forecasting Division
Human Effectiveness Directorate
711th Human Performance Wing
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE ***(DD-MM-YYYY)***	**2. REPORT TYPE**	**3. DATES COVERED** ***(From - To)***
30-Nov-2011	Final	June 2010 –November 2011

4. TITLE AND SUBTITLE
Analyst Performance Measures Volume III
Information Quality Tools for Persistent Surveillance Data Sets

5a. CONTRACT NUMBER FA8650-09-6939
5b. GRANT NUMBER NA
5c. PROGRAM ELEMENT NUMBER 62202F

6. AUTHOR(S)
John Talburt
Brian Tsou
Mariofanna Milanova
Serhan Dagtas
Mihail Tudoreanu

5d. PROJECT NUMBER 7184
5e. TASK NUMBER 006
5f. WORK UNIT NUMBER H06J (7184X18W)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Arkansas at Little Rock
2801 South University Avenue
Little Rock, AR 72204

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Forecasting Division
Human-Analyst Augmentation Branch
Wright-Patterson AFB OH 45433-7022

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER (S)
AFRL-RH-WP-TR-2012-0114

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A: Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
88ABW/PA cleared 24 September 2012 as 88ABW-2012-5092.

14. ABSTRACT
The Air Force desires a comprehensive vehicle to identify and address requirements for information quality tools and techniques to support defensive and offensive operations research in the layered sensing domain. As use of remote sensors in the Air and Space domains increases, the value of the sensor datasets must be maximized, and assurances established that the product outcomes meet the application requirements. As multiple sensors are combined into layered sensing systems, this increases the need to understand not only the quality and fitness for using the individual sensor data streams, but also how to assess the quality and value of the aggregate data.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Brian Tsou
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	SAR	147	**19b. TELEPHONE NUMBER** ***(include area code)*** (937) 255-8896

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.18

THIS PAGE LEFT INTENTIONALLY BLANK.

TABLE OF CONTENTS

Section **Page**

LIST OF FIGURES

LIST OF TABLES

1.0 OVERALL SUMMARY

The Air Force desires a comprehensive vehicle to identify and address requirements for information quality tools and techniques that will support defensive and offensive operations research in the layered sensing domain. As use of remote sensors in the Air and Space domains increases, the value of the sensor datasets must be maximized and assurances established that the product outcomes meet the application requirements. As multiple sensors are combined into layered sensing systems, this increases the need to understand not only the quality and fitness for using the individual sensor data streams, but also how to assess the quality and value of the aggregate data.

The scope of this task order is to develop metrics that assess the quality and effectiveness of persistent surveillance data sets. The project also explores the use of three dimensional (3D) visualization in rendering layered data sets and experiments with the integration of textual information. In addition, integrating processing of data available from multiple types of sensors (such as in a Smart Environment) has been explored, and experiments have been done to support data fusion for multiple sensors.

The work is divided into three tracks being worked concurrently by three different research teams. Each research team consists of one Principal Investigator and two Graduate Students.

Track One: Identify Information Requirements for the Layered Sensor Domain focuses on the identification of Information Quality (IQ) metrics as related to analysis of video data streams. Analysis of current research identifies some work done in this area. Within this track, further experiments resulted in the identification of new Information Quality metrics for video data streams.

We demonstrated the weighted, new objective quality metrics on an intuitive example. We used traffic video data containing 23 frames from a ground sensor camera and then distorted an original, reference video using three different processes: Blurring, Salt and Pepper Noise and Joint Photographic Experts Group (JPEG) compression. Each process has three distortion amounts.

Our results show that our metrics are more realistic and correlated than existing metrics. In the future, we will develop a subjective quality assessment to validate our metrics with human subjective perception.

Two novel image quality metrics, Saliency-Based Structural Similarity Index (S-SSIM) and Saliency-Based Visual Information Fidelity (S-VIF) in pixel domain, were also developed. The metrics are based on frequency-tuned, salient region detection and computationally inexpensive. Experiments show that the proposed metrics match with Human Visual System better than Structural Similarity Index (SSIM) and Visual Information Fidelity (VIF) in pixel domain.

During the summer of 2009, one initial experiment was developed related to video quality and tracking moving objects. The main objectives of the experiment were to develop a new moving objects' tracking algorithm and to investigate the quality.

Track Two: Prototype the Utilization of Interactive 3D Information Visualization in the Layered Sensor Domain focuses on using interactive 3D visualization to improve the quality of information in the Layered Sensor domain. The research considered national Aeronautics Space Administration (NASA) World Wind—an application mostly used in two dimensional (2D) settings and successfully ported it for use in a 3D environment known as a Cave Automatic Virtual Environment (CAVE).

Track 2 successfully created new techniques for improving the quality (IQ) of the Layered Sensors Domain. The techniques are modular in nature and can be combined with each other or with other World Wind layers.

The presentation of data in a CAVE can let the user feel immersed in it and better understand 3D relationships. A demonstration at Tec^Edge required only a few hours to set up a portable immersive system. The software included both the CAVE port and Twitter on World Wind.

The integration of Twitter data with Geographic Information System (GIS) data promises to increase the quality of both data sources. Twitter has very good timeliness, but may suffer from accuracy, noise, and believability, while GIS data is largely correct, but slightly outdated. Combining the two sources provides an overall increase in quality because one can draw on the other's strengths.

Track Three: Visual Rendering and Display of Text & IQ Metrics – Smart Environment postulates that integrated processing of data available from multiple types of sensors can benefit a variety of decision-making processes. An information processing scheme is prototyped that offers data fusion for multiple sensors such as temperature sensors or motion detectors and visual sensors such as security cameras.

Track 3 demonstrated ways to use the Bayesian data fusion technique in a smart environment with a heterogeneous, inter-dependent set of sensors. This was done by generating statistically independent inputs for the Bayesian fusion model and demonstrating the effect through a simulation tool.

The *Dempster-Shafer theory* is considered to be a generalization of the Bayesian theory of subjective probability. Dempster-Shafer allows us to "base degrees of belief for one question on probabilities for a related question" [6]. One of the most important advantages of the Dempster- Shafer theory is that it does not associate probabilities to questions of interest as Bayesian methods do. Instead, the belief for one question is based on probabilities for a related question; therefore, the Dempster-Shafer theory can effectively model uncertainty.

Detailed information on each track's research and results are presented within this report.

2.0 INTRODUCTION TO TRACK ONE - IDENTIFY INFORMATION - REQUIREMENTS FOR THE LAYERED SENSOR DOMAIN

Motivation: The industry's need for accurate and consistent objective video metrics has become more critical with new digital video applications and services such as Internet video, surveillance, mobile broadcasting and Internet Protocol Television (IPTV).

The study focus was on the fundamental needs of emergency responders to communicate and share information in an effective and timely manner in Layered Sensor Doman.

***Challenge*:** Video quality metrics have been proposed in order to predict the human visual perception and to achieve high correlation with the human perception.

2.1 Objectives

- Identify and document the information requirements for the Layered Sensors Domain and to investigate the state-of-the-art research and information quality methods and techniques for the Layered Sensors Domain
- Develop information quality metrics appropriate for layered sensor data streams, investigate the relationship between information quality metric values and applications outcomes, and develop strategies for embedding information quality metadata "tags" into sensor datasets

2.2 Research Development Plan

- Perform a literature search for publications in scientific and technical research journals, conference proceedings, and other venue in order to document and build on existing knowledge in Video and Multimedia IQ.
- Investigate and understand information requirements of actors involved in the Layered Sensors Domain.
- Explore the problems and issues related to the integration of multiple sensor data streams.
- Develop information quality metrics appropriate for layered sensor data streams, investigate the relationship between information quality metric values and applications outcomes, and develop strategies for embedding information quality metadata "tags" into sensor datasets.
- Develop new metrics and demonstrate how the video quality is measured for video records where the task is tracking moving objects

2.3 Methods, Assumptions, and Procedures

2.3.1. Assumptions

Quality of Experiences (QoE) has become a term commonly used to describe application – and user-oriented quality of video and multimedia services. In [13], Winkler and Mohandus listed some of the numerous factors contributing to quality of multimedia data:

- Individual interests of the viewer, such as favorite sources of information, which determine the level and focus of attention

- Quality expectations of the viewer; for example, film screened in a cinema versus a short clip watched on a mobile device
- Video experience of the viewer (once you have seen high-definition content, it's hard to go back)
- Display type (size, resolution, brightness, contrast, color, and response time)
- Viewing setup and conditions, such as viewing distance or ambient/exterior light
- Quality of synchronization of different sensor information
- Interaction with the service or display device (e.g. remote control)

Most of the existing video quality metrics only account for a small subset of the factors listed above and focus on measuring the visual fidelity of the video in terms of the distortion introduced by various processing steps (mainly compression and transition). The following challenging issues still remain unsolved:

- Video systems are complex and consist of many components, including capture and display hardware, converters, multiplexers, codes, streamers, routers, and switches.
- Digital multimedia contents are subject to a wide variety of distortions during transmission, acquisition, processing, compression, storage and reproduction any of which may result in degradation of visual quality.
- These distortions depend on economics and/or physical limitations of the devices.
- Visual perception is even more complex. We need to understand how people perceive video and its quality.

2.3.2. Methods and Procedures

Video Quality Assessment (VQA) methods fall into two categories: 1) subjective assessment by humans and 2) objective assessment by algorithms.

Subjective image quality experiments are classical statistical measurements how humans pensive the image quality. Subjective measures are determined by Mean Opinion Score (MOS) which relies on human perception. The mathematical tools for subjective assessment of image quality are well defined, but still there remain certain practical aspects how to design efficient experiments. While subjective assessment is the ultimate judge of image quality, it is time consuming and cannot be implemented in real time quality score. This is the main reason to motivate development of algorithms which predict subjective image quality measure accurately. In [1] how "well" an algorithm performs is defined by how well it correlates with human perception of quality

Objective quality metrics are algorithms designed to characterize the quality of image and predict viewer opinion. Different types of objective metrics exist as illustrated in [1], [2]. They are based on mathematical measurements which are practical to apply without need of human observers. Objective quality metrics can be classified into 3 metrics: Full Reference (FR), Reduced Reference (RR) and No Reference (NR). All these metrics are based on the availability of original non-distorted reference image which will be compared with the corresponding distorted image. In a FR case, reference image information is available; in a RR case, partial information of reference image is known and no information about the reference image is available in the NR case.

In the image processing community more than 50 years Mean Squared Error (MSE) are being used as quasi–standard fidelity metrics. The MSE still continue to be widely used as a signal fidelity measure, but at the same time there are recent studies to developed more advanced signal fidelity measures, especially in applications where perceptual criteria might be relevant.

The approaches in metrics design can be classified in two groups: 1) a visual modeling approach and 2) engineering approach [13]. We have developed algorithms using the engineering approach.

- **Objective Image/Video Quality Metrics Test**
    - **Test One: MSE**

MSE is widely used as it is parameter free, computationally simple and mathematically convenient in the context of optimization. It also represents image energy measure that energy is preserved after any orthogonal linear transformation, such as the Fourier transform. However, MSE does not fit precisely with the perceived visual quality. Distorted images with the same MSE may have different visibility. [3] [4]

Consider two images $x = \{x_i \mid i = 1,2,..., N\}$ and $y = \{y_i \mid i = 1,2,..., N\}$ where N is the number of pixels and x_i and y_i are the ith pixels of the images of x and y, respectively; the MSE between these two images is:

$$\mathrm{MSE}(x, y) = \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2 \qquad (1)$$

- **Test Two: SSIM**

Consider two images $x = \{x_i \mid i = 1,2,..., N\}$ and $y = \{y_i \mid i = 1,2,..., N\}$ where N is the number of pixels and x_i and y_i are the ith pixels of the images of x and y, respectively. SSIM- $\mathrm{SSIM}(x, y)$ combines three comparison components, namely luminance-$l(x, y)$, contrast-$c(x, y)$ and structure-$s(x, y)$ [5]:

$$\mathrm{SSIM}(x, y) = f(l(x, y), c(x, y), s(x, y)) \qquad (2)$$

Luminance, contrast and structure comparisons are defined as follows:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad C_1 = (K_1L)^2$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad C_2 = (K_2L)^2 \qquad (3)$$

$$s(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3},\quad C_3=\frac{C_2}{2}$$

where μ_x, μ_y, σ_x, σ_y and σ_{xy} are means of x and y, variances of x and y and correlation coefficient between x and y. K_1 and K_2 are scalar constants that $K_1, K_2 << 1$ and L is the dynamic range of the pixel values. Finally, SSIM index yields to:

$$\mathrm{SSIM}(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \tag{4}$$

- **Test Three: VIF in Pixel Domain**

VIF index relates image fidelity to the mutual information between the test and the reference images using source and distortion models and as well as human visual system model. It is given as [6]:

$$\mathrm{VIF}=\frac{\sum_{j=1}^{S}\sum_{i=1}^{M_j} I(C_{i,j};F_{i,j})}{\sum_{j=1}^{S}\sum_{i=1}^{M_j} I(C_{i,j};E_{i,j})} \tag{5}$$

$I(C_{i,j};F_{i,j})$ and $I(C_{i,j};E_{i,j})$ represent the information perceived by the human observer from a particular sub band in the reference and the test images respectively. C is a block vector from a given location in the reference image, E is the perception of block C by a human observer from reference image, which can be represented as $E = C + N$, where *n* is additive noise. F is the perception of block C by a human observer from test image, which can be represented as $E = D + N$. D is the block vector from the test image given as $D=GC+V$ where G and V are the blur and noise distortions, respectively. S denotes the number of all sub-bands and M_j is the number of blocks at j th sub-band.

- **New Objective Quality Metrics**
  - **Weighted Objective Quality Metric When the Task is Tracing Moving Objects in Video**

In human visual system, the importance of a visual event should increase with the information content, and decrease with the perceptual uncertainty [7], we incorporated foreground mask (see Appendix 1) as weighting function into the MSE and SSIM metrics to measure the motion feature of the moving car. At a time MSE is $\mathrm{MSE}(x,y,t)$ and SSIM is $\mathrm{SSIM}(x,y,t)$.

The weighting function is:

$$w(x,y,t)= I(x,y,t)-median\{I(x,y,t-i\} > \tau \tag{6}$$

We define weighted MSE as wMSE and weighted SSIM as wSSIM as follows:

$$wMSE = \frac{\sum_x \sum_y \sum_t w(x,y,t)\mathrm{MSE}(x,y,t)}{\sum_x \sum_y \sum_t w(x,y,t)} \tag{7}$$

$$wSSIM = \frac{\sum_x \sum_y \sum_t w(x,y,t)\mathrm{SSIM}(x,y,t)}{\sum_x \sum_y \sum_t w(x,y,t)} \tag{8}$$

- **Attention-Based Weighted Objective Quality Metric**

The important aspect in video quality evolution is the fact that people only focus on certain regions of interest in the video. On the base of our previous work on model of attention, two new metrics are developed.

S-SSIM and S-VIF in Pixel Domain: In human visual system, the importance of a visual event should increase with the information content, and decrease with the perceptual uncertainty [8]. We incorporated saliency map as weighting function into the SSIM and VIF indexes. So saliency factors can be instated into the quality metrics. The weighting function is:

$$w(x,y) = \left\| I_\mu - I_{w_{hc}}(x,y) \right\| \tag{9}$$

We define saliency-based SSIM as S-SSIM and saliency-based VIF as S-VIF as follows:

$$\begin{aligned} \mathrm{S\text{-}SSIM} &= \frac{\sum_x \sum_y w(x,y)\mathrm{SSIM}(x,y)}{\sum_x \sum_y w(x,y)} \\ \mathrm{S\text{-}VIF} &= \frac{\sum_x \sum_y w(C,F)\mathrm{VIF}(C,F)}{\sum_x \sum_y w(C,F)} \end{aligned} \tag{10}$$

$$wSSIM = \frac{\sum_x \sum_y \sum_t w(x,y,t)\mathrm{SSIM}(x,y,t)}{\sum_x \sum_y \sum_t w(x,y,t)} \tag{11}$$

SSIM and VIF in pixel domain mainly focus on local information and do not take global saliency features into consideration [9]. Figure 1 shows an example case that SSIM and VIF in pixel domain fail. It is easy to see that the quality of images in Figure 1(d) and Figure 1(f) are better than that of Figure 1(c) and Figure 1(e). Even though the amounts of distortion effects are greater in Figure 1(c) and Figure 1(d), SSIM and VIF in pixel domain give incorrect results.

(a) Reference Image

(b) Saliency Map of the Reference Image

(c) Distorted Image with Higher Amount of Gaussian Noise Applied to Attended and Less- Attended Locations

(d) Distorted Image with Less Amount of Gaussian Noise Applied to Only Less-Attended Locations

(e) Distorted Image with Higher Amount of Blurring Effect Applied to Attended and Less- Attended Locations

(f) Distorted Image with Less Amount of Blurring Effect Applied to Only Less-Attended Locations

Figure 1: An Example Case that SSIM and VIF in Pixel Domain Fail

As shown in Table 1, S-SSIM and S-VIF in pixel domain scores are more realistic.

Table 1: Scores of SSIM, S-SSIM, VIF and S-VIF in Pixel Domains for Images in Figure 1

	SSIM	S-SSIM	VIF in pixel	S-VIF in pixel
Figure 1(c)	0.5976	0.8319	0.6886	0.2196
Figure 1(d)	0.724	0.5772	0.7774	0.1449
Figure 1(e)	0.3851	0.865	0.6751	0.3136
Figure 1(f)	0.4463	0.6452	0.9336	0.2436

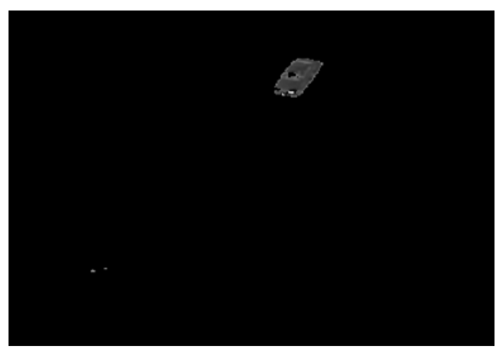
2.4 Results and Discussion

2.4.1. Results Implementing wMSE and wSSIM

We demonstrated the weighted *new* objective quality metrics on an intuitive example. We used a traffic video data containing 23 frames from a ground sensor camera. We distorted the original reference video generated from three different types of processing: Blurring, Salt and Pepper Noise and JPEG compression. Each process has also three distortion amounts.

We presented a novel objective quality assessment metric. In proposed metrics, moving objects from video sequences are particularly considered as visually important content. Background subtraction based on approximate median filter is used for tracking the moving objects. Then foreground masks are computed from the absolute difference of estimated background and input frame. Existing metrics, MSE and SSIM, are modified by the weighting factors of the foreground masks. We applied our approach to a traffic video data from a ground sensor.

(a) Estimated Background



(b) Foreground Mask

Figure 2: Estimated Background and Foreground Mask

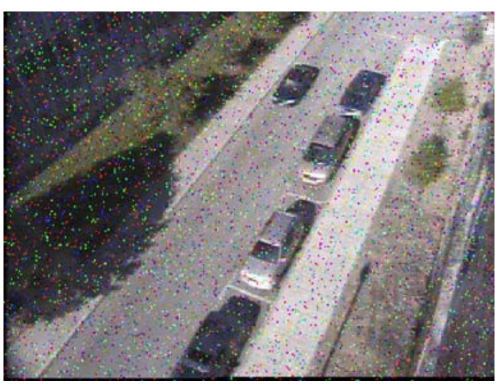
Our results show that our metrics are more realistic and correlated than existing metrics. In the future, we will develop a subjective quality assessment to validate our metrics with human subjective perception.

Table 2: Distortion Processing and Amount

Distortion Type	Distortion 1	Distortion 2	Distortion 3
Blurring	fil. size=6, std. dev = 6	fil. size=8, std. dev = 8	fil. size=10, std. dev = 10
Salt and Pepper	d (noise density) = 0.01	d (noise density) = 0.03	d (noise density) = 0.05
JPEG Compression	compression = 50%	compression = 70%	compression = 90%

(a) Sample Reference Frame

(b) Blurred of Size 10 with Standard Deviation 10

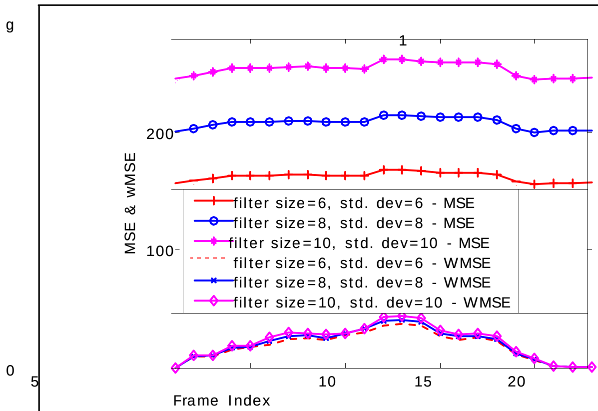


(c) Salt and Pepper Noise with Noise Of 0.05

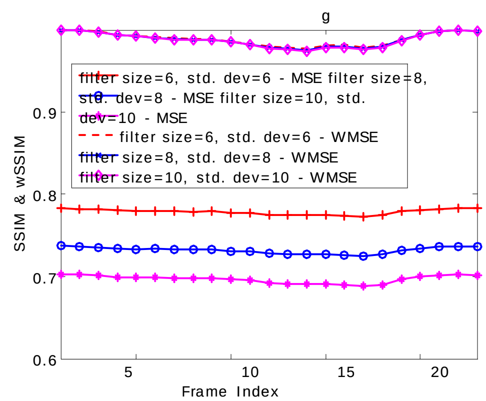
(d) JPEG Compression with 90%

Figure 3: A Sample Frame Image from the Video Data and Associated Distortions

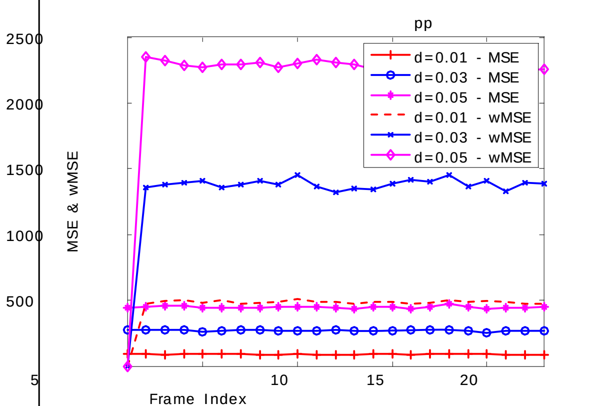
Figure 4 shows the results of objective VQA. As shown in the figures, weighted metrics are more realistic and correlated with human perception. For instance, since there is no moving car in the first frame, MSE and SSIM give wrong scores while weighted metrics give 0.0 and 1.0, respectively, as they give importance to only moving content. Similarly, in other frames, wMSE values are less than those of MSE, and wSSIM values are greater than those of SSIM. This is because visually important content such as the moving car is considered more by wMSE and wSSIM.



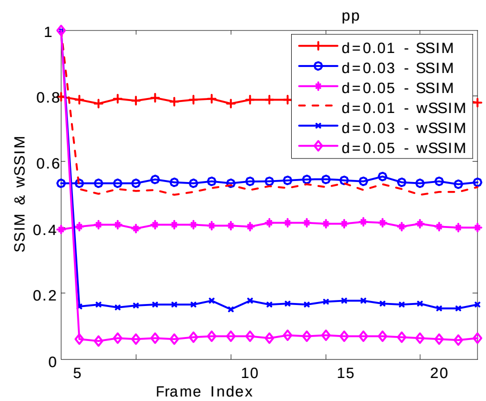
(a) MSE And MSE with Proposed Weighting Method for Blurring Distortion



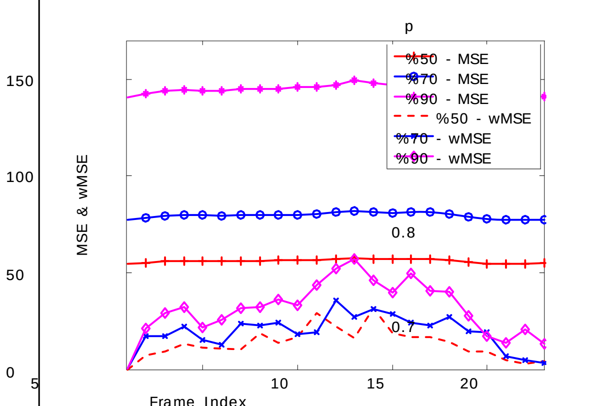
(b) SSIM And SSIM with Proposed Weighting Method for Blurring Distortion



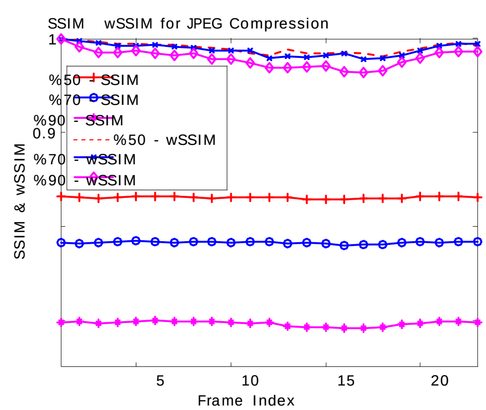
(c) MSE And MSE with Proposed Weighting Method for Salt & Pepper Effect



(d) SSIM And SSIM with Proposed Weighting Method for Salt & Pepper Effect



e) MSE and MSE with Proposed Weighting Method for JPEG Compression



f) SSIM And SSIM with Proposed Weighting Method for JPEG Compression

Figure 4: Objective VQA Plots on a Test Video Containing 23 Frames

The results using wMSE and wSSIM were presented in a paper to the "Recent Advances in Signal Processing, Robotics and Automation" conference in March 2010.

2.4.2. Results Implementing S-SSIM and S-VIF

We validated our approach using two image databases as test bed: These databases contain subjective scores for each image. First is the Image and Vision Computing (IVC) database [11] consisting of 10 reference images with 235 distorted images (JPEG, JPEG2000, Locally Adaptive Resolution (LAR)-coded and blurred). Second is the Laboratory for Image & Video Engineering (LIVE) Image Database hosted at the University of Texas at Austin [12] consisting of 29 original images and 460 distorted images (227 JPEG2000 images and 233 JPEG images). Non-linear regression analysis has been performed to fit the data. The Pearson correlation coefficient is used to measure the association between subjective and objective scores.

Figures 5 and 6 show the results for IVC and LIVE databases, respectively. Each sample point represents the subjective/objective scores of one test image. The y axis in the figures denote the subjective scores in the databases. The x axis denotes the predicted quality of images after a nonlinear regression toward four objective scores, which are SSIM, S-SSIM, VIF and S-VIF in pixel domains, respectively.

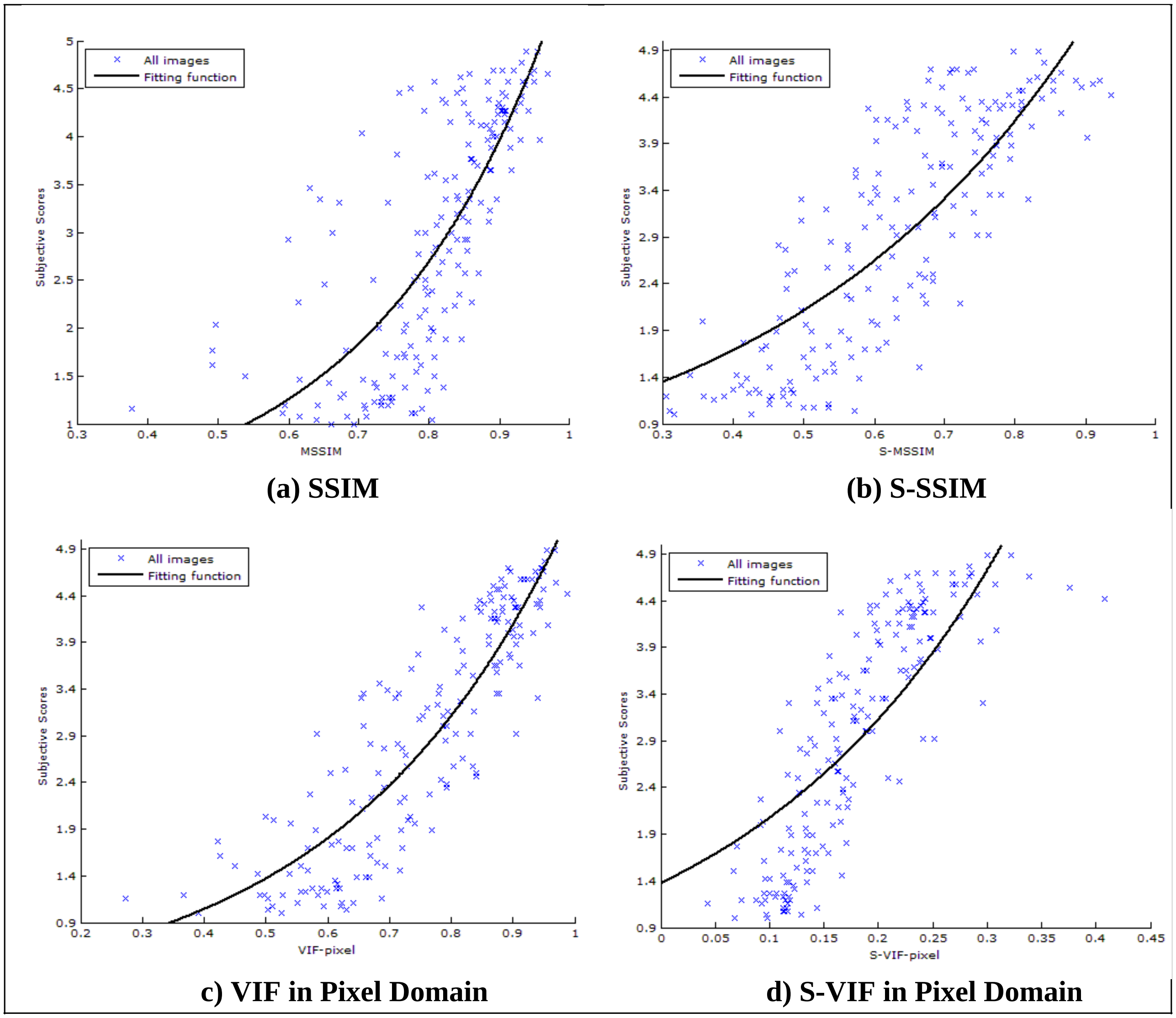
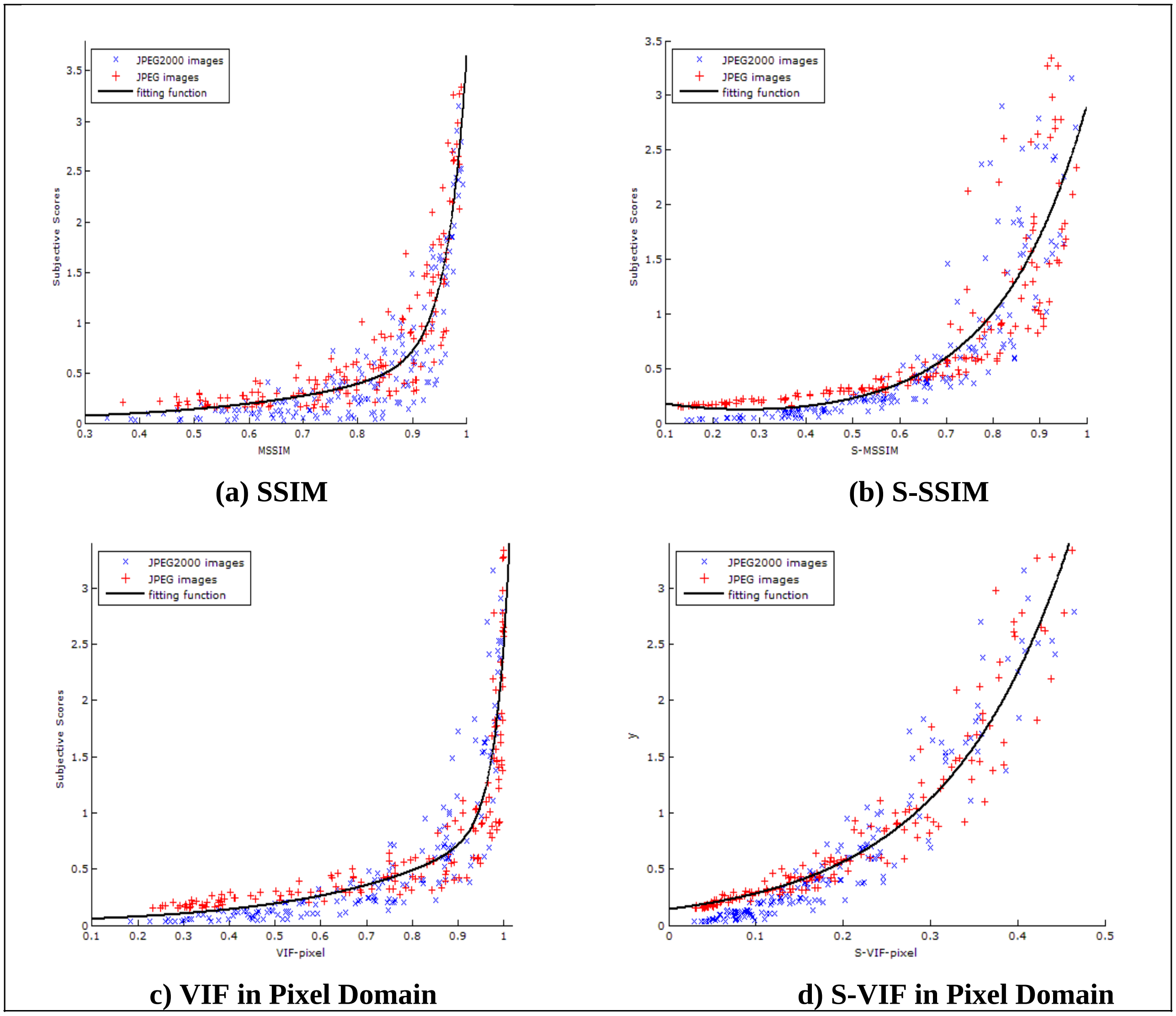


Figure 5: Scatter Plots of Subjective/Objective Scores on IVC Database



(a) SSIM **(b) S-SSIM**

c) VIF in Pixel Domain **d) S-VIF in Pixel Domain**

Figure 6: Scatter Plots of Subjective/Objective Scores on LIVE Database
(Red Points and Blue Points Denote JPEG and JPEG2000 Images, Respectively)

The Pearson correlation coefficient varying from -1 to 1 is widely used to measure the association between two variables. High absolute values mean that the two variables being evaluated have high correlation. As shown in Table 3, our technique is more correlated with human subjective perception.

Table 3: Pearson Correlation Coefficients

	SSIM	S-SSIM	VIF-Pixel	S-VIF-Pixel
IVC - All Images	0.7047	**0.8261**	0.8435	**0.8715**
LIVE - JPEG&JPEG2000 Images	0.6823	**0.7475**	0.7126	**0.9083**

Two novel image quality metrics, S-SSIM and S-VIF in pixel domain were developed. The metrics are based on frequency-tuned salient region detection and computationally inexpensive. salient region detection captures full resolution saliency maps exploiting the color and luminance features of the images. Saliency maps are then set as weighting functions and incorporated into SSIM and VIF in pixel domain. The approach has been validated using two image databases: 1) IVC Image database consisting of 10 reference images with 235 distorted images (JPEG, JPEG2000, LAR-coded and blurred) and LIVE Image Database consisting of 29 original images and 460 distorted images (227 JPEG2000 images and 233 JPEG images.). Experiments show that the proposed metrics match with Human Visual System better than SSIM and VIF in pixel domain.

The results using S-SSIM and S-VIF were presented in a paper entitled, "Image Quality Assessment Based on Salient Region Detection." (See Appendices)

2.4.3. Results Implementing Image Quality of Tracking Moving Object

During the summer of 2009, one initial experiment was developed related to video quality and Tracking Moving Object. The main objectives of the experiment are to develop new moving objects tracking algorithm and to investigate the quality changing the following Image Quality Dimensions:

- Resolution (the physical area that a single pixel covers)
- Noise
- Brightness
- Contract
- Saturation
- Gamma Connection

These results were presented in a paper entitled, "Image Quality of Tracking Moving Object." (See Appendices).

2.5 Conclusions

As mentioned earlier, the existing quality metrics have many shortcomings:

- They measure video degradation. In surveillance applications, video fidelity even considering the characteristics of the human visual system is clearly not a quality benchmark in such applications.
- Only few metrics attempt to model the focus of attention and consider it for computing overall video quality.
- There is a need to develop hybrid metrics.

Future studies will be conducted to develop quality metrics when different sensor data are synchronized.

2.6 Recommendations

It is interesting to demonstrate how the image quality is measured for different regions in an image. It is obvious that different regions in the image may not stand the same importance. Visual importance has been explored in the context of visual saliency [14], fixation calculation

[16]. In [15], one experiment to record the gaze coordinates corresponding to the human eye movements and the Gaze-Attentive Fixation Finding Engine (GAFFE) was proposed. In [16], the researchers are using GAFFE to find points of potential visual importance and one algorithm for fixation-based and quality- based weighting was developed. The region-of-interest-based image quality assessment still remains unexplored.

It is interesting to work on Hybrid Metrics [13].

2.7 Introduction Year Two Work

The second year of work on Information Quality Tools for Persistent Surveillance Data Sets extends research done during the first year. Track 1 research was extended to focus on the development and assessment of a new perception-based image/video quality metrics using CIELAB Color Space called the Structure of Color Structural Similarity Metric (C-SSIM).

2.7.1. Color Model and Visual Perception

Human color vision is trichromatic, consisting of three cone signals transformed into three channels: red-green opponent channel; a blue-yellow channel; and a luminance channel. Mr. Billock and Mr. Tsou provided an interesting study about human color vision [6]. Why these color channels? Mr. Billock and Mr. Tsou show that many aspects of visual color perception can be explained by assuming that the three color channels form the axes of a vector space. Such a color space defined by dimension of lightness or luminescence (L) and two color components “a” and “b” is called a LAB. The French Commission Internationale de L'éclairage LAB (CIELAB) describes all of the colors visible to the human eye. Due to the fact that color values are linearized with respect to perceptual color differences, a measured change in color value can cause the same relative change in the visual properties of an image. In human vision system, the relationship between objectively derived image features and a subjective image’s quality score obtained from a grader can lead us to new discoveries. In the CIELAB color model, the “L” stands for luminescence, “a” is the magenta contrast, and finally “b” is named as the yellow contrast.

Davis at al [7] presents more information regarding classifying the most important color features related to image quality.

Feature A: It is indicated as mean intensity for each channel. Increasing or decreasing of exposure of mean intensity can affect image quality.

Feature B: For each color channel, skewness can be defined as a measure of asymmetry or symmetry of intensity distributions.

Feature C: For each color channel, kurtosis shows a measure of pixel intensities with respect to normal distribution.

A full characterization of the luminance channel for any image can be obtained by utilizing these features. In this study, Davis concludes that the “b“ or yellow contrast occurs five times in the top ten weighted features and the weight contribution of this feature is 48% of the overall weight. The contribution of ”L” or luminescence is 17%, and lastly the “a” or magenta contrast is 35%.

2.7.2. The RGB to LAB Transform

First the *RGB* tristimulus values are transformed into device independent *XYZ* tristimulus values. It is common practice to use a device-independent conversion that maps white in the chromaticity diagram to white in *RGB* space and vice versa [8].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.5141 & 0.3239 & 0.1604 \\ 0.2651 & 0.6702 & 0.0641 \\ 0.0241 & 0.1228 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

(6)

The device independent *XYZ* value are then converted to *LMS* space by

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

(7)

Redurman et al [9] obtained a color space, namely Lab. It efficiently reduces the correlation between the *LMS* axes. Redurman et al are using the following simple transform to decorrelate the axes in the logarithmic *LMS* space:

$$\begin{bmatrix} L \\ a \\ b \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}$$

(8)

2.7.3. Structure of Color Structural Similarity Metric

The color structural similarity metric $C\text{-}SSIM_{color}$ is defined as:

$$C - SSIM_{color} = \sqrt{w_L}(S_L)^2 + \sqrt{w_a(S_a})^2 + \sqrt{w_b}(S_b)^2$$

(9)

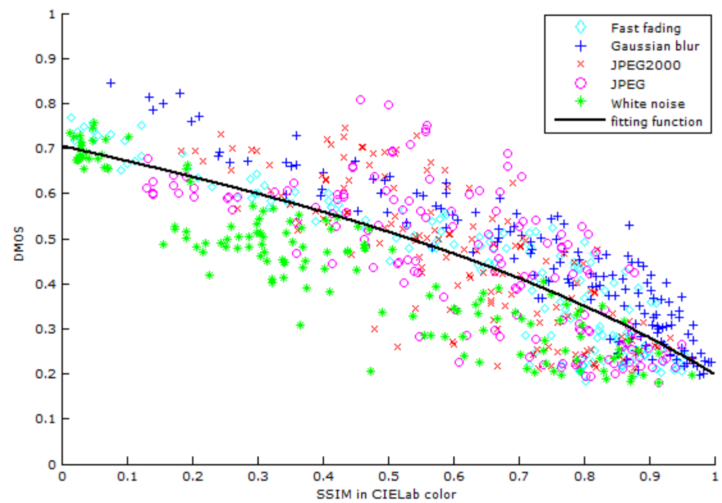
where S_L, S_a and S_b are respectively the structural similarity factors given by Eq.(3), computed for each of the individual LAB color channels and w_L, w_a, w_b are the corresponding weights credited to the perceived distortions in each of these channels.

2.8 Results and Discussion

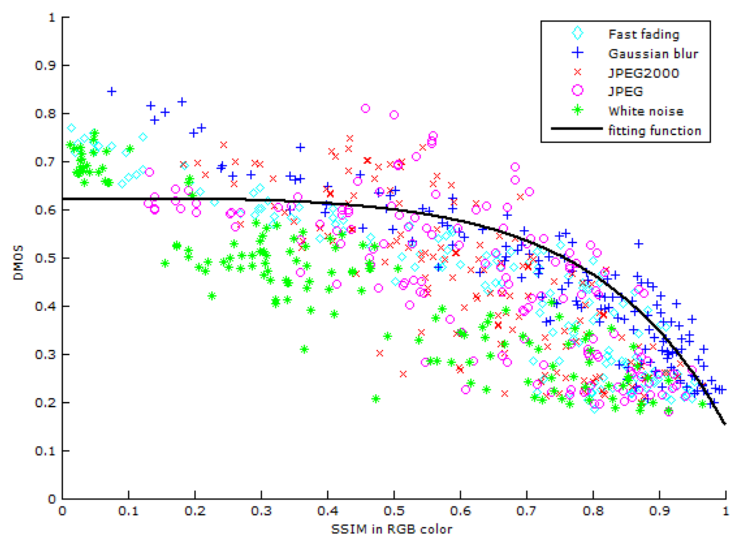
2.8.1. Results Implementing Perception-Based Image Video Quality Metrics Using CIELAB Color Space

In order to test and validate our proposed quality metric, the LIVE Image Database Release 2 [10] was chosen as a test bed. The LIVE Image Database consists of 29 high-resolution 24 bits/pixel color reference images (typically 768 x 512) and their distorted images (982 images) under five distortion types: JPEG2000, JPEG, white noise, Gaussian blur, and bit errors. Each distorted image has a computed Difference Mean Opinion Score (DMOS) ranging from 1 to 100. JPEG2000 images were generated using various bit rates. White noise images were obtained using White Gaussian noise. Gaussian kernel was used to create Gaussian blurred images. Fast- fading Rayleigh channel model was utilized to generate transmission errors in JPEG2000 bit stream.

To assess the correlation between our color metric and human visual perception, we performed an extensive experiment using the LIVE Image Database. In this experiment, SSIM value was computed for each distorted image in RGB and CIELAB color spaces respectively. Obtained data was illustrated as scatter plots in Figure 1. Non-linear regression analysis was performed to fit data. Each sample point in the scatter plots has corresponding SSIM and DMOS values. DMOS values are normalized to 1 and represented in the y-axis, whereas SSIM values are located in the x-axis of scatter plots. As shown in Figure 1, as DMOS increases, SSIM value decreases meaning that they are in inverse proportion. DMOS versus SSIM in CIELAB color space shows inverse proportion as seen Figure 1a, whereas DMOS versus SSIM in RGB color space poses near a quadratic relation in Figure 1b. Therefore, SSIM in CIELAB, namely proposed Color-SSIM, correlates well with DMOS under various distortion types.



(a)



(b)

Figure 7: Scatter Plot of DMOS versus (a) SSIM in CIELAB Color Space (b) SSIM in RGB Color Space

The results using CIELAB color space were presented in a paper entitled, “Perception-Based Image/Video Quality Metric Using CIELAB Color Space.” Authors were Sertan Kaya, Travis Bennett, Mariofanna Milanova, John Talburt, Brian Tsou, Marina Altynova, and Hongyan Xu , SPIE Proceedings Vol. 8019” (See Appendices).

2.8.2. Results Implementing Neural Network

The LIVE database [10] is selected as a test bed to perform experiments and validate our approach. The LIVE image database contains 29 high-resolution 24 bits/pixel color reference images and theirs distorted images under five distortion types: JPEG2000, JPEG, White noise, Gaussian blur, and bit errors. Here, we chose 50 distorted images from each distortion type. Each distorted image has a computed DMOS ranging from 1 to 100. JPEG2000 images were generated using various bit rates. White noise images were obtained using White Gaussian noise. Gaussian kernel was used to create Gaussian blurred images. Fast-fading Rayleigh channel model was utilized to generate transmission errors in JPEG2000 bit stream.

The experiment consists of two major steps, namely, training and testing. The training section encompasses three aspects as follows; creating feature vectors, obtaining target vectors and designing neural network architecture. To create feature vector, we divide an image into grids such that 8 x 8 sliding window can scan through entire image. In this 8x8 window, statistical features such as mean, standard deviation and covariance are computed for each original and distorted image pair for a batch process. This process is repeated for each four pairs of original and distorted images among five pairs. The reason is that four pairs of images (original and distorted) are used for training and one pair is considered for testing purposes. By doing so, the feature vectors are generated for a type of image. This is for only one batch training process. This process is repeated for 50 image pairs for each distortion type. Obtaining target vectors are essentially based on subjective score, namely DMOS, carried out by human observers. The DMOS value is already provided for each distorted image given in the LIVE database set. To be able to obtain the DMOS value corresponding each 8x8 window, we use mean weighed technique as follows: We calculate mean of each window and entire image. The DMOS corresponds to mean of the entire image and computed DMOS is assigned each window based on their weighed mean. With this fashion, target vectors are generated corresponding feature vectors. As we mentioned in the previous section, selected neural network architecture is multilayer feed forward neural network. The back propagation is chosen as a learning algorithm for proposed framework. The number of hidden layers is composed of three and each hidden layer consists of six neurons. The logistic sigmoid activation function is used in the hidden layers and the linear activation function is employed in the output layer. Using feature vectors and target vectors under determined neural network architecture, training process is achieved which gives a net which is saved for testing section. The Matrix Laboratory (MATLAB) training interface screen shot is depicted in Figure 2 for training section.

In the testing section, the left one pair image among five pairs for the same type of images is used to yield input vectors with the same method explained above, as the MATLAB testing interface screen shot illustrated in Figure 3. After the input vectors are fed into the neural network system, the ultimate goal is obtain DMOS. The output of neural network would be the predicted DMOS. This process is repeated for all 250 images under various distortion types with their original images. With this procedure, obtained data is depicted in the scatter plot in Figure 4. Non-linear regression analysis was performed to fit data. Each sample point in the scatter plots has corresponding DMOS and neural network output values. DMOS values are represented in the y-axis, whereas predicted DMOS values are located in the x-axis of scatter plot. As shown in Figure 4, as DMOS values increase, predicted DMOS values increase meaning that they are in proportion.

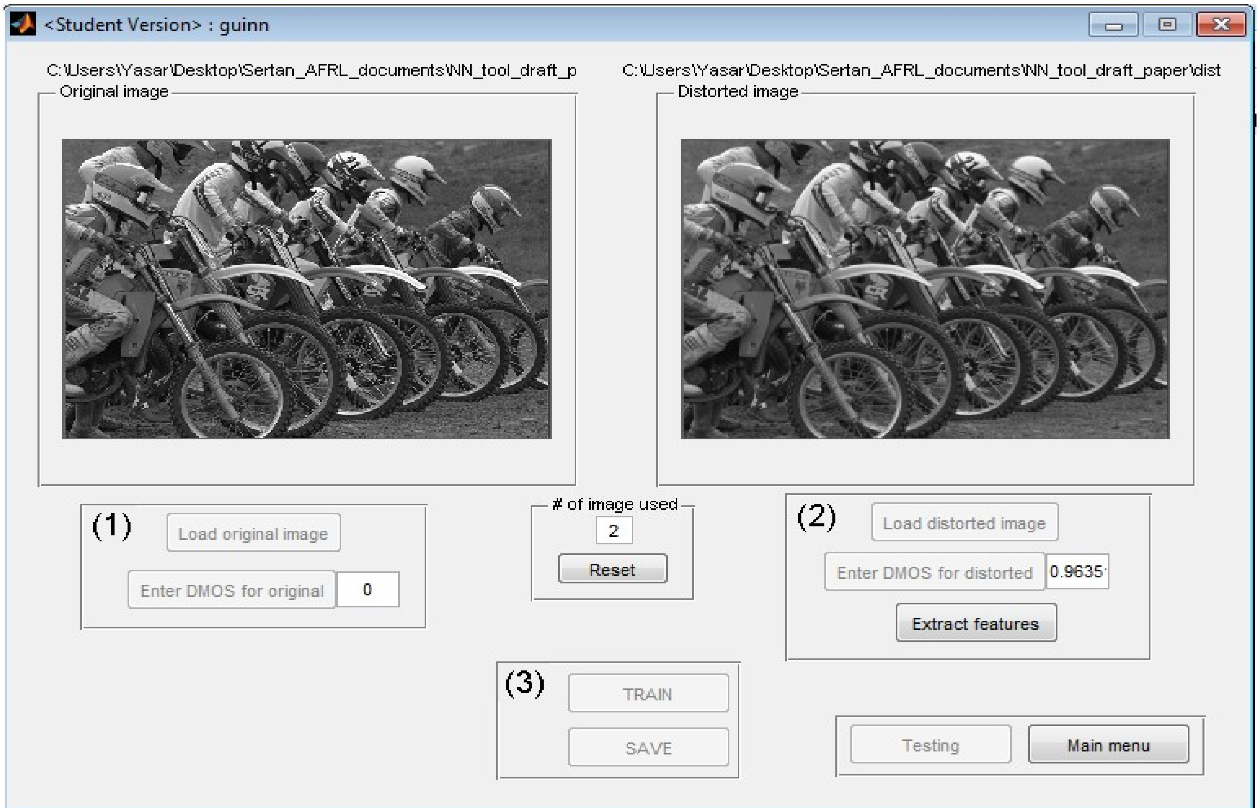


Figure 8: The MATLAB Training Interface

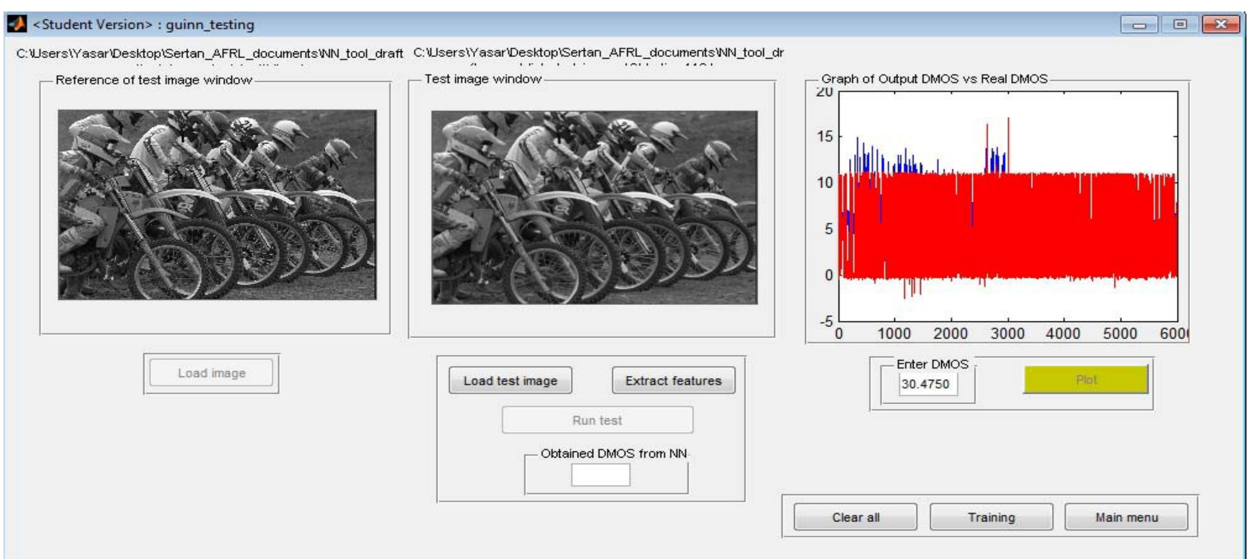


Figure 9: The MATLAB Testing Interface

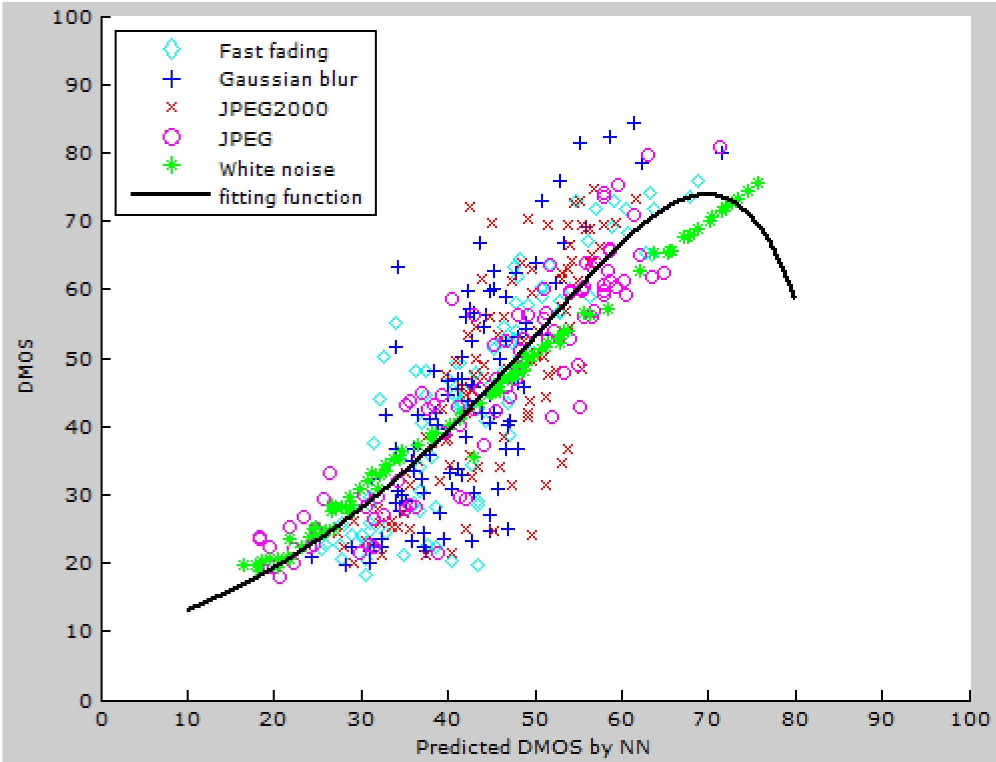


Figure 10: Scatter Plot of DMOS versus Predicted DMOS by NN

The results using Neural Network for image quality assessment were presented in a paper entitled, “Subjective Image Quality Prediction Based on Neural Network.” (See Appendices).

2.8.3. Results Implementing Video Segmentation and Annotation Tool

During Fall 2011 and Spring 2011, a new tool was developed related to video annotation:

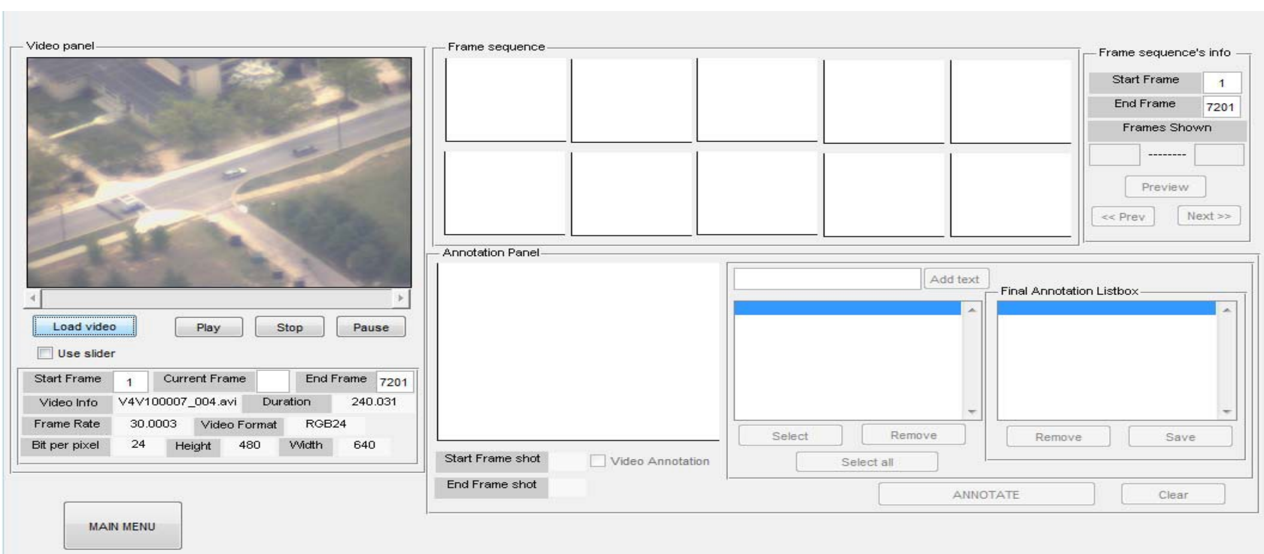


Figure 11: (Step 1) Upload a Video by Clicking Load Button

(Slide Features Lets User Scroll Through Video)

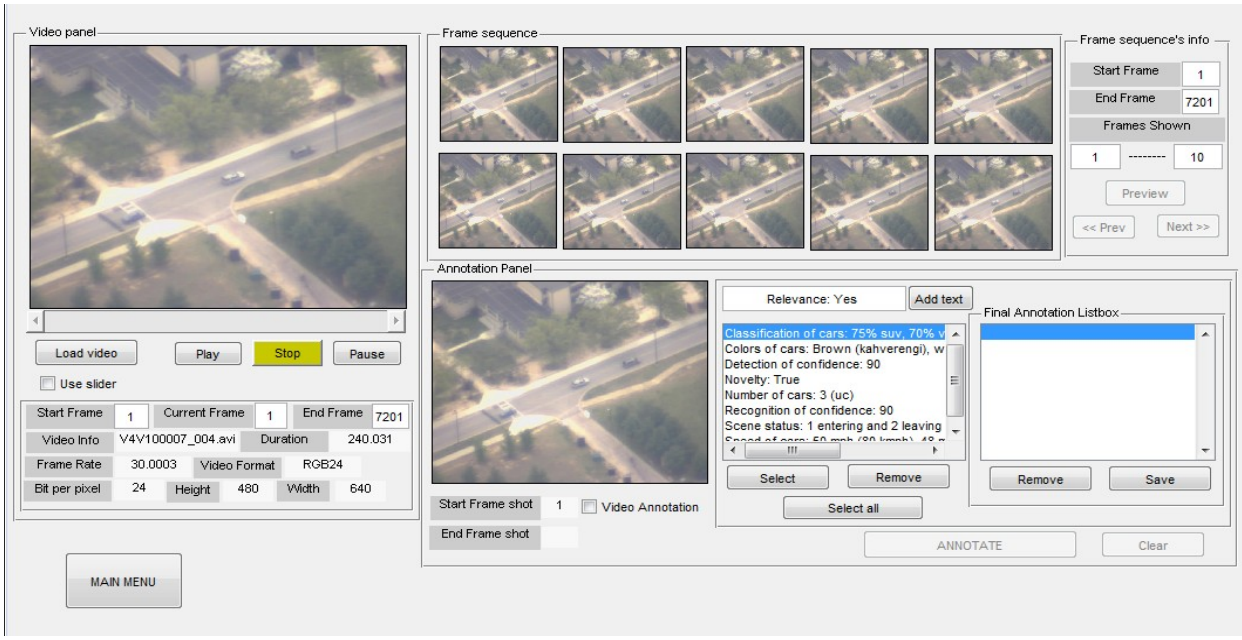


Figure 12: (Step 2) Lets the User Select Any Frame Sequences in the Video

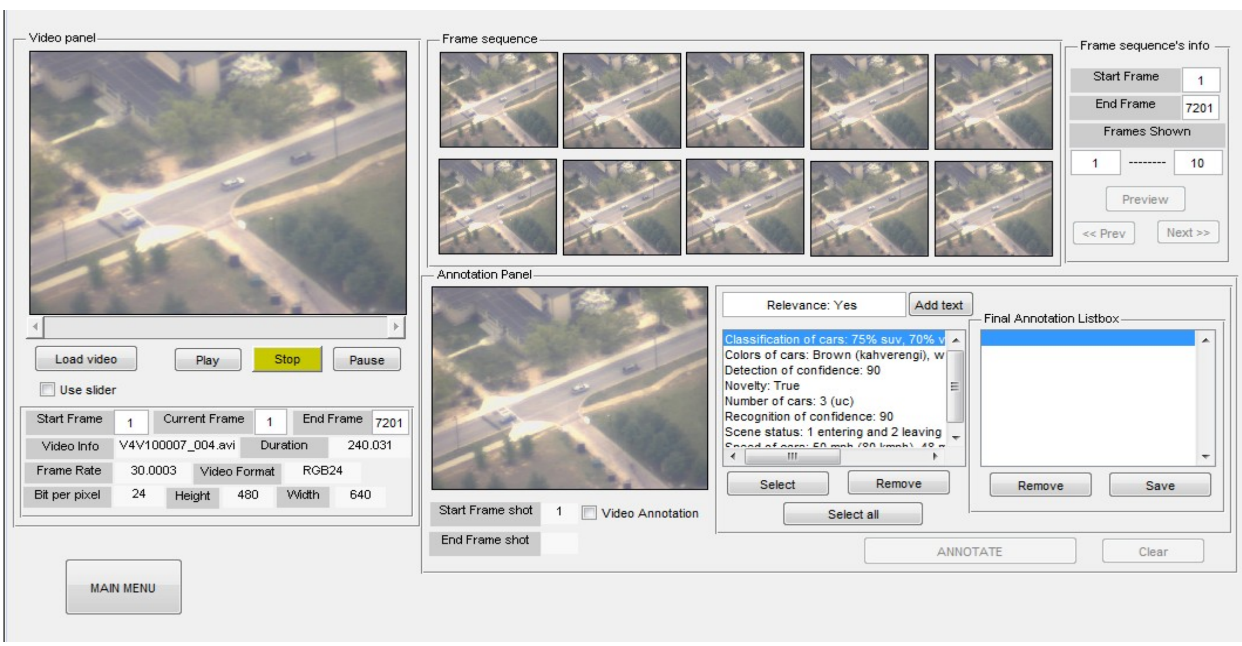
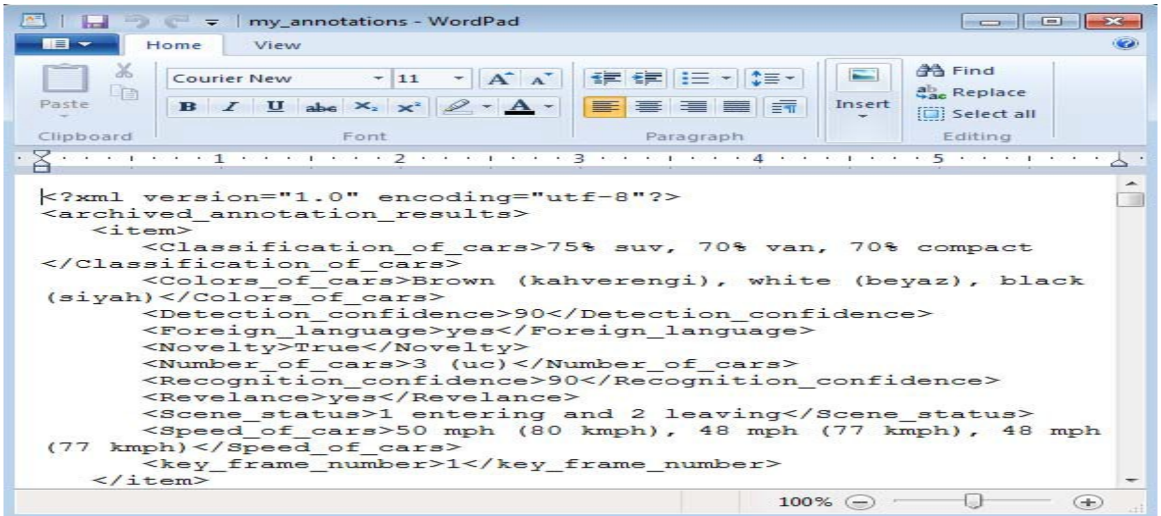


Figure 13: (Step 3) Image Annotation




```
<?xml version="1.0" encoding="utf-8"?>
<archived_annotation_results>
   <item>
      <Classification_of_cars>75% suv, 70% van, 70% compact
</Classification_of_cars>
      <Colors_of_cars>Brown (kahverengi), white (beyaz), black
(siyah)</Colors_of_cars>
      <Detection_confidence>90</Detection_confidence>
      <Foreign_language>yes</Foreign_language>
      <Novelty>True</Novelty>
      <Number_of_cars>3 (uc)</Number_of_cars>
      <Recognition_confidence>90</Recognition_confidence>
      <Revelance>yes</Revelance>
      <Scene_status>1 entering and 2 leaving</Scene_status>
      <Speed_of_cars>50 mph (80 kmph), 48 mph (77 kmph), 48 mph
(77 kmph)</Speed_of_cars>
      <key_frame_number>1</key_frame_number>
   </item>
```


Figure 14: (Step 4) Saved Annotations Corresponding Key Frames in an XML File

These results were presented in a paper entitled, "Volkan H. Bagci, Mariofanna G. Milanova, Roumen Kountchev, Roumiana Kountcheva, Vladimir Todorov: Object and Scene Recognition Using Color Descriptors and Adaptive Color KLT. HCI (12) 2011: 355-363" (See Appendices).

2.9 Conclusions

We explored color image quality assessment's problem by introducing a novel metric, Color-SSIM derived from CIELAB color model. Our motivation was the high sensitivity of SSIM to a wide variety of geometric transformation of image in spatial domain. We applied SSIM to each channel of CIELAB color space separately and then put them into a weighed vector mean. We validated the applicability of our new metric by extensive testing with the LIVE Image Dataset Release 2. Experimental results demonstrated that Color-SSIM correlates better than SSIM in RGB color space with DMOS

3.0 INTRODUCTION TO TRACK TWO - PROTOTYPE THE UTILIZATION OF INTERACTIVE 3D INFORMATION VISUALIZATION IN THE LAYERED SENSOR DOMAIN

Track Two focused on the utilization of interactive, 3D information visualization to improve quality of information in the Layered Sensor Domain. The purpose of our work was to improve the situational awareness of a user by finding novel ways of presenting existing sensor data and by combining Internet-based, non-spatial data into the same view.

Track Two considered NASA World Wind, a GIS application that is mostly used in 2D settings, and successfully ported it for use in an immersive, 3D environment commonly known as a CAVE. The Internet data consisted of live and recorded Twitter short messages.

Our avenues of research included:

- Design of new data processing mechanisms to allow seamless integration of GIS and Twitter
- Development of several visualization software prototypes including CAVE and regular desktop solutions as well as GIS and non-GIS centered
- Demonstration of the prototypes and techniques both at the University of Arkansas at Little Rock (UALR) CAVE and at Tec^Edge in a mobile immersive settings
- Evaluating the effectiveness of the mechanisms, interaction, and visualization tools
- Publication in peer-reviewed venues

3.1 Methods, Assumptions and Procedures

Track Two research will be presented in three main components: (1) porting a version of NASA World Wind to immersive systems, which are devices with multiple, stereo displays that make use of tracking devices for the hand and head of the user; (2) Twitter visualization in order to provide situational awareness; and (3) 3D models of building inclusion in NASA World Wind.

3.1.1. World Wind for CAVE

Track Two successfully built a prototype of World Wind that is capable to run in any immersive display. We presented live software demonstrations at Tec^Edge in an ad-hoc, two screen immersive environment. People can fly through the world via a Wanda device, which tracks user's hand motion. A portable Wanda was brought to Tec^Edge from UALR for this demonstration.

The prototype was also tested and used in the UALR CAVE, which is another type of immersive environment. The CAVE has three stereo displays, each measuring 10' x 10', and are placed in a half-cube arrangement. User can walk through this cube and a head tracker is attached to stereo glasses to sense user's movement. A Wanda device is used for interaction. The CAVE is powered by a cluster of three computers.

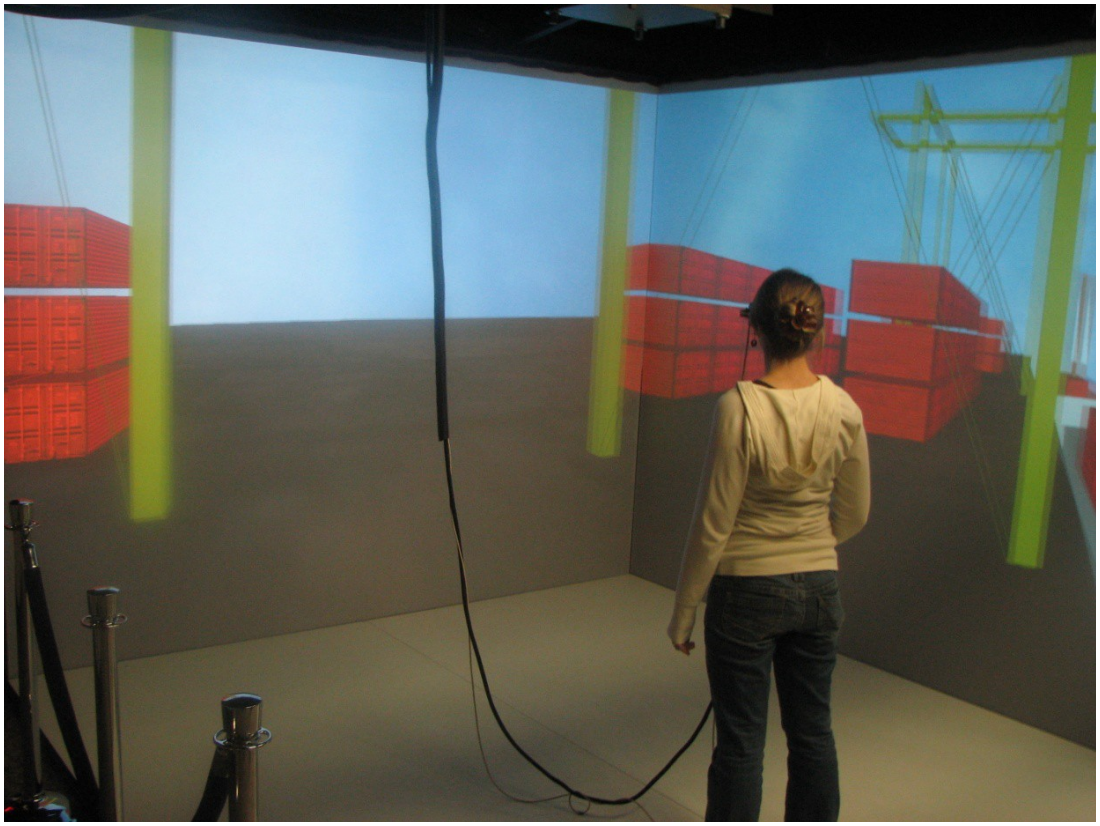
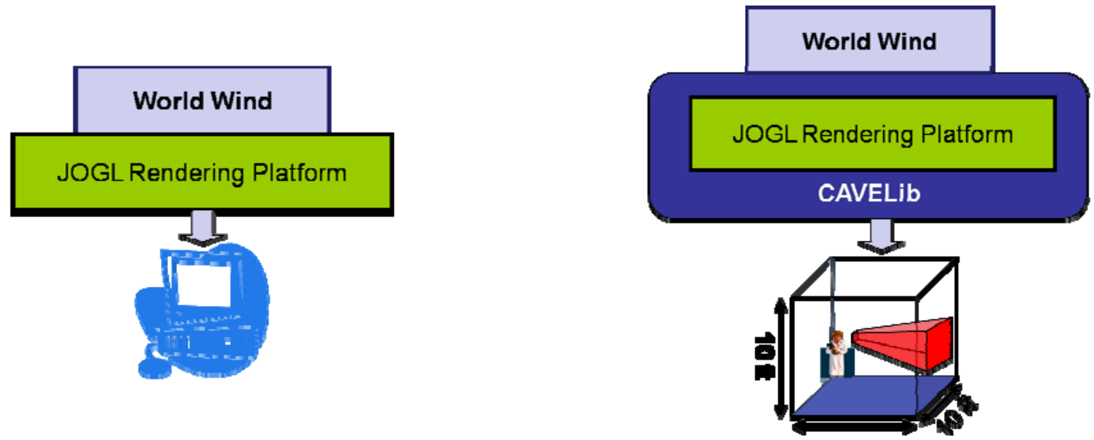


Figure 15: UALR CAVE System with Three Stereo Displays and Head and Hand Tracking Devices

The immersive World Wind prototype was designed to be able to run on any number of possible immersive display configurations. It is based on a commercial Application Programming Interface (API) named CAVELib, which provides a lightweight platform for any 3D application. CAVELib is commercial software purchased by UALR, and it is meant to work with C/C++ programs.

The main challenge was to integrate CAVELib, a C/C++ application, into the Tec^Edge version of World Wind, which is entirely written in Java. Our work focused on changing the underlying platform of World Wind, named Java Open Graphical Language (JOGL) to accept commands and configurations from CAVELib. Development work included a thorough understanding of JOGL architecture and an appropriate integration solution with CAVELib.



(a) Original World Wind Configuration **(b) Immersive World Wind Configuration**

Figure 16: Differences between Rendering Platforms of World Wind Desktop and Immersive World Wind

The second task was to create the appropriate view projection and interaction mechanisms for a user in the immersive environment. The perspective into the virtual world is more extensive than on a regular desktop that uses a flat screen and a mouse. To this end, we developed a series of Java classes to handle the scene, projection, and interaction in the new World Wind. These classes were integrated seamlessly into World Wind and can be simply chosen by modifying the existing built-in configuration file.

Interaction with the immersive World Wind was designed to allow users to use their hand and a virtual 3D wand to "fly" through the world. Users can also use a combination of button presses and wrist movement to rotate the virtual world around them on all three coordinate axes.

3.1.2. Twitter Visualizations

Two main avenues were explored for adding Twitter data to the layered sensors domain: show the spatial distribution of Twitter data and build visualizations of the Twitter discourse trends. The goal was to explore the mechanism and benefits of including non-GIS, Internet data into the layer sensor, and to improve the IQ of both sources.

***Twitter on World Wind*:** The main thrust of this part was to produce a Twitter layer capable of displaying the geographical distribution of Twitter messages. The volume of users broadcasting tweets makes Twitter a source of vast amounts of various kinds of information. However, this magnitude of information created and broadcasted with little or no restriction other than the amount of characters contained in a tweet presents both opportunities and challenges to data mining efforts. It presents opportunities in the sense that it provides real time information about individuals, the public's perspective on issues as well as current news events. Some of the problems that we perceive with this information are potential amount of "noise" created by wrong or unverifiable information.

Although Tom Anderson, a social media market researcher, described Twitter as a “Babylon of Spam” [1], it can also be argued that it is also a source of valuable (current) information. For example, Twitter played an important role in broadcasting information from within Iran during the Iranian election crises of 2009 [2]. Moreover, security forces used Twitter as a source of real time information during the Mumbai terrorist attack in 2008 [3], fire department and weather monitoring organizations also provide updates to the public via Twitter [4].

The most challenging part of the process was the vastly different format of the two types of data source. Tweets are unstructured, non-spatial text, and GIS information is well organized and spatial. We devised a technique of simultaneously processing and analyzing textual and map information in order to create a common visualization. The interaction requirements of our techniques are virtually the same as typical GIS exploration. Furthermore, we examine an information visualization method that helps users easily digest the integrated information

Our approach exploits the logical linkage between the non-spatial textual data and the GIS information. The exact origin on Earth of the text may never be known, but the meaning refers to specific places. Our research produced techniques both for discovering the linkage and for visualizing the distribution of important textual keywords on the map. We believe that for most tasks and applications, the logical relationship between GIS and text will prove important, especially because it may be the only type of relationship available to infer.

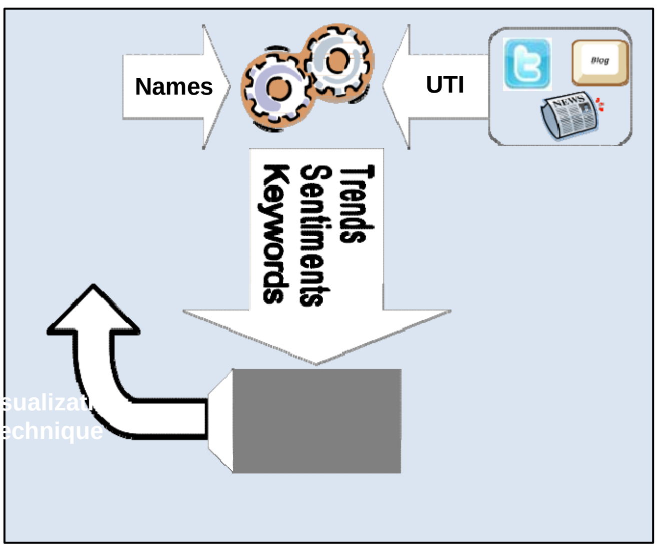


Figure 17: GIS Information is Extracted and Used to Assign Location to Unstructured Text

Next we describe our approach to assigning spatial positions for non-spatial data, in particular unstructured textual data. Three topics are covered: (1) the reasoning behind our choice of linkage between the two categories of information, (2) extraction of relevant information from both sources, and (3) the placement of textual data as a word cloud visualization on the map.

The Reasoning Behind the Choice of Linkage Between Two Categories of Information: By nature, unstructured text does not have standard geographical coordinates. Hence, the geospatial representation of textual information is not typically a straight forward endeavor. Therefore, a critical task in integrating textual information with geospatial information is determining the best algorithm for joining different categories of data. The algorithm depends on factors such as the intended use of the final information product, the scope of the textual information, and the available data manipulation technologies. For example, to develop an application that uses the Institute for Electrical and Electronic Engineers (IEEE) publications database to determine the concentration of information visualization researchers in the US. The name of organizations to which artificial intelligence researchers are affiliated would be a good attribute to join research publications and GIS information. Generally speaking, the question may be more complicated in the absence of a clear structure of the document and a lack of a spatial attribute. For such unstructured data, a more logical or semantic search can be performed to determine the textual data for which geo-coordinates can be inferred.

Our proof of concept is focused on geospatially representing Twitter information on World Wind in an application that can be used for example by first responders to increase their awareness of a theatre of operation. There is no direct widely available attribute that can be used to link the two categories of data. Although the geo-location feature of tweets (contains the geo-coordinates of the tweet origin) can be used to join both information categories, there are a couple of quickly visible drawbacks to this approach.

Firstly, only a fraction of tweets have values for this feature (according to eWeek.com [5] only 0.23%), which would render the vast majority of Twitter data un-joinable. The second drawback is more specific to the intended use of the application. For the described usage of our application, the information contained in the tweet body provides more information about a certain location than the origin of the tweet. For instance, a tweet might be sent from a hotel room at point A, but it concerns events happening in the home city of B. Hence, the logical link between GIS and information contained in the tweet better serves the intended purpose of our application than a join based on the tweet origin.

Considering that our main aim for introducing Twitter information into World Wind is to get a feel for the buzz about geographical locations contained in the Twitter chatter, we employed place/location names as the basis for exploring the logical linkage. This approach gives us the ability to assign to tweets geographical coordinates of location(s) whose name they contain. In general, any location for which geo-coordinates exist can be used (e.g., state, counties, cities, landmarks, organizations and street) as can databases with various public servants names such as mayors or governors. We used the location names available to the PlaceName layer of World Wind. The layer contains the names of continents, countries, towns, and cities as well as their geo-coordinates.

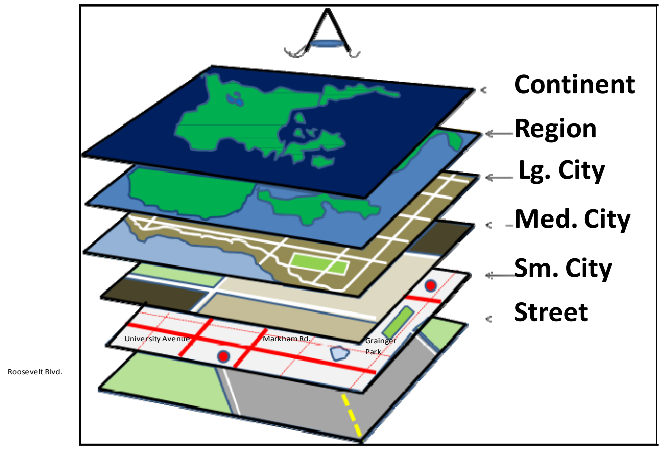


Figure 18: Information Layers According to Level of Details

Extraction of Relevant Information from Both Sources: This section describes the process of extracting the required information from both the GIS and textual sources. Systematically extracting and requiring information not only makes the application efficient, but also makes it user friendly.

Given the large geographical distances that can be covered in a relatively short time on a GIS application like World Wind, a lot of place names must be extracted from the PlaceName layer and used as a query argument to extract tweets of interest from Twitter. We approached this task by using some of the existing mechanisms in World Wind to decide what information to request/extract from its databases. In particular, we determined from the World Wind interface the geographical area in view of the user. Various place names are associated with that area at different levels of details as shown in Figure 18.

Our approach is to extract names from multiple levels of detail, even if World Wind does not currently render that level (because the user may be at a high altitude). The place names collected from World Wind are used in the Twitter query to obtain the tweets that contain those names. Effectively, the user queries Twitter simply by flying from one place to the other on the globe. Consequently, the number of places flown over is directly proportional to the number of queries performed on Twitter. This approach has the general effect that the user can explore multiple information sources while expending only the same usual amount of effort required for maneuvering on World Wind.

Although World Wind, like many other GIS applications, manages the level of details displayed for a geographical area with respect to altitude (altitude of user's view), this level of detail depends on several factors that might not contain enough information for extracting data from the UTI source. For example, if the GIS shows information at the country level but the UTI does not contain country names, then it would be impossible to extract information from the UTI based on that particular string (country name). We used information with finer details beyond that displayed on the map for querying Twitter. For example, if the user is at the country level, we dig deeper to obtain further information about the states or cities in view of the map.

This level of detail allows us to extract more data from Twitter (and presumably more information) about the location. After this information is analyzed, the results are aggregated with respect to the geographical space and presented to the user. This also has the advantage that information from Twitter would have already been requested (and possibly obtained and analyzed) before the user reaches lower altitudes (see Figures 19 and 20 for different altitudes/level of detail)

In order for the user to access the knowledge contained in the UTI without having to read through several lines of text (especially when s/he is pressed for time), the UTI should be analyzed and presented in an easily interpretable format to the user. The type(s) of analyses to be carried out depends on the intended use of the final product, the type of UTI, as well as the technology. We performed three key analyses on the set of tweets obtained from each query result namely: keywords analysis, sentiment analysis, and trend analysis

We used the keyword "analysis" to get an overview of the main topics of the discourse in the query results for a particular place. The analysis was performed by identifying the most frequent words other than English language stop words 1 and the query argument 2 contained in the tweets. The result of the analysis contains each word in the result set frequency. We attempted to use the sentiment analysis to capture the emotion/mood (happy, sad, and panic) expressed in the tweets. Depending on the use of this application, the mood about a place can trigger different actions. For example, detection of a panic emotion may prompt the police to increase physical presence. We used a pre-compiled list of words that signal each of the three moods that we analyzed. The mood of the query result is decided by determining which of the buckets of pre-compiled emotion words is most represented in the entire keywords set (i.e., the bucket with the most number of words that can be found in the keywords set). The trend analysis is simply a way to create some persistence in the data analysis. It depicts fluctuations (if any) in the Twitter mood and most frequent keywords in a place over time. This is important as it can help in identifying unusual patterns in the information obtained for a place.

The Placement of Textual Data as a Word Cloud Visualization on the Map**:** Two main types of strategies for rendering keywords in a spatial environment were developed, and they differ in whether single points or entire areas are considered as anchors for displaying Twitter data. The visualization strategies balance the two competing goals of presenting a large number of Twitter keywords and of creating a clear, easy to understand view.

- **Point-Based Strategies**

Point-based strategies start by directly assigning the location of each query (e.g., city, street name) to the Twitter keywords that correspond to that query. This simple approach results in keywords being displayed on top of each other and, if visible, on the query name (location). This problem can be corrected by artificially spreading out the keywords and even by changing their font size as needed. Either random alteration of the keyword placement or some regular, geometric pattern can be implemented.

---

[1] The language of the stop words is dependent on the language of the text.
[2] The query argument (place name) was not included in the analysis because, it does not provide any additional information and it is contained in the all tweets in the query result.

Point-based strategies display very precisely the association of queries to keywords, but aggregation occurs through spatial placement, and may hide some important keywords in favor of relatively unimportant ones. This "visual aggregation" occurs on the map in the sense that for a given area, such as a state, all keywords about the cities in that state are shown next to each other. A lack of explicit aggregation may lead to a situation in which keywords with low frequency may be displayed while some relatively frequent keywords are not visible. Consider the case of two neighboring cities, one small and the other large. The small city may be able to display all its associated keywords, even those that only occur once or twice in tweets. The large city may have a large number of keywords all with a high frequency, yet due to space constraints, only a fraction of those keywords are shown. This is exactly the case of unimportant keywords being visible (around the small city) while frequent terms are cut out (around the large city).

- **Area-Based Strategies**

The second type of strategy, area-based, can emphasize overall aggregation of keywords. All the Twitter terms that fall within a given area are considered together and only the most important ones are displayed. The size of the area is dependent on how far the user is from Earth's surface, and we implemented this using standard GIS tiles. Keywords are placed around the area in such a way to avoid overlap, and they are regarded as part of the area rather than belonging to a point on the map. This provides a larger real-estate for placement, and can alleviate repetition of keywords.

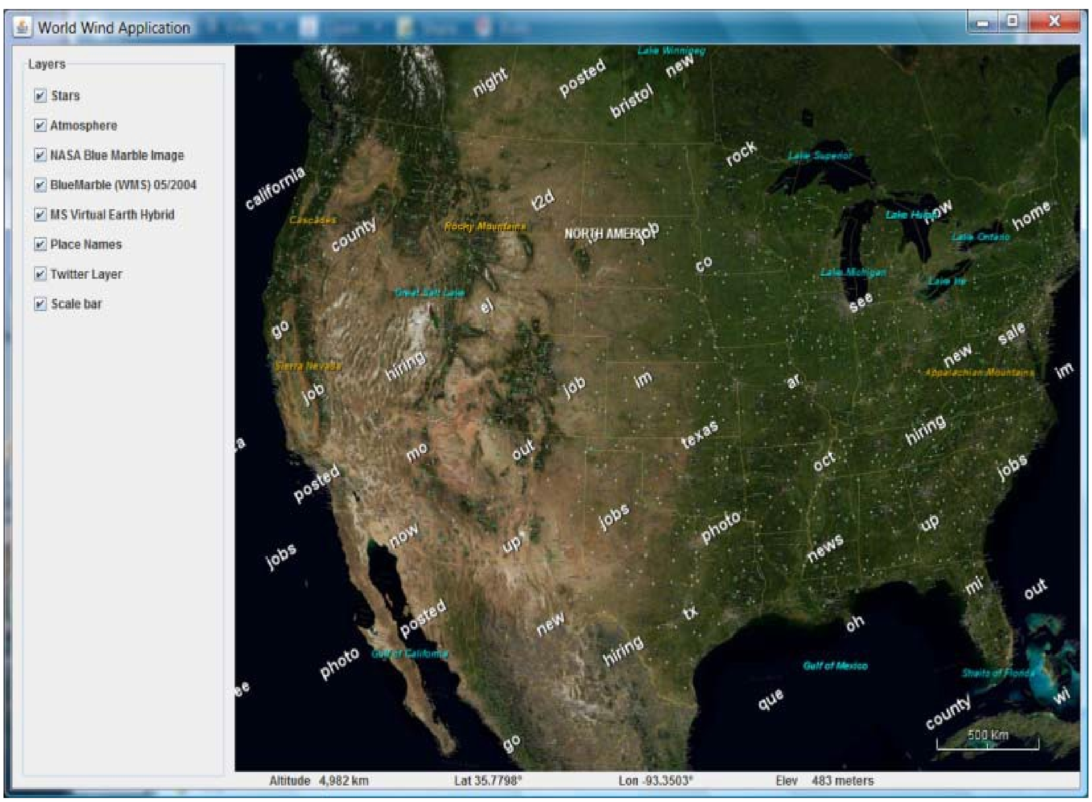


Figure19: Twitter Keywords Placement Approximation at the Country Level View on World Wind

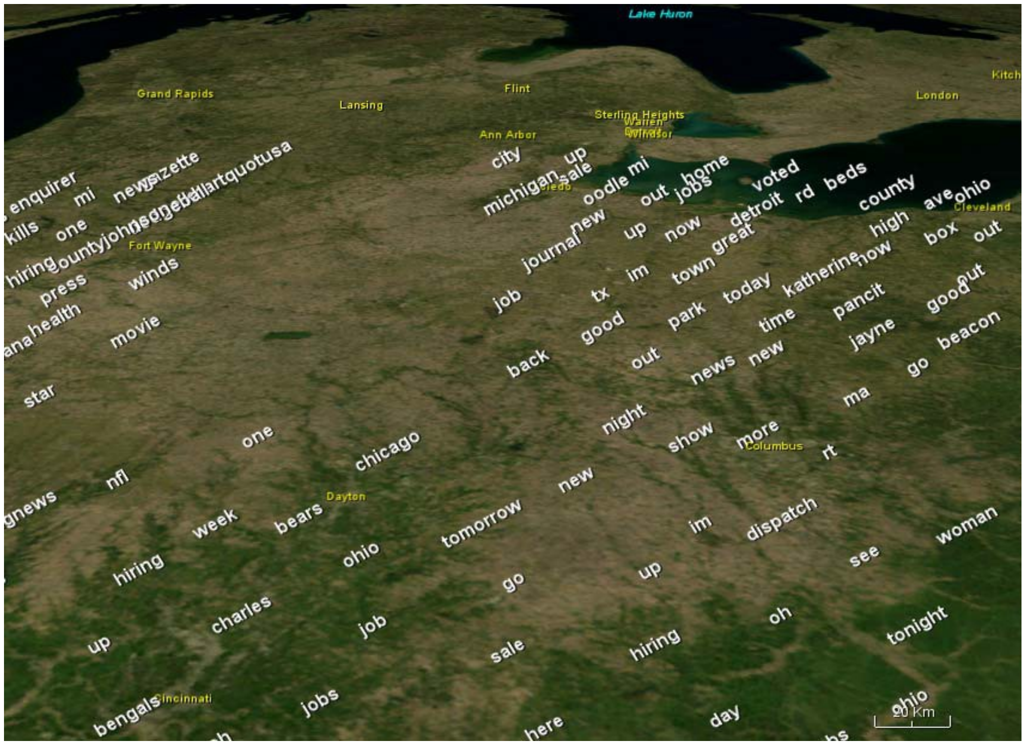


Figure 20: Regional View of Twitter Data

There are three techniques for spreading out Twitter terms in an area:

- Random placement
- Geometric pattern, such as concentric circles or grids
- Weighted-average placement of keywords

The first two techniques are computationally inexpensive, with the second being able to also convey the relative importance of the terms by, for example, placing the most important term in the center and using a pre-defined ordered placement after that. The third technique requires the use of either forces or virtual “bungee cords” to pull a keyword in its final position. The anchor points for the forces or cords are the location of the queries associated with the keyword (e.g., the cities in which the keyword is tweeted). This approach is similar to MonkEllipse [6], and it will result in popular/widespread keywords appearing in the center of the area because they are “pulled” in multiple directions towards most places in that area. The averaged position may also lead to overlap, and requires an extra overlap-reducing step in which keywords lying on top of each other are spread around.

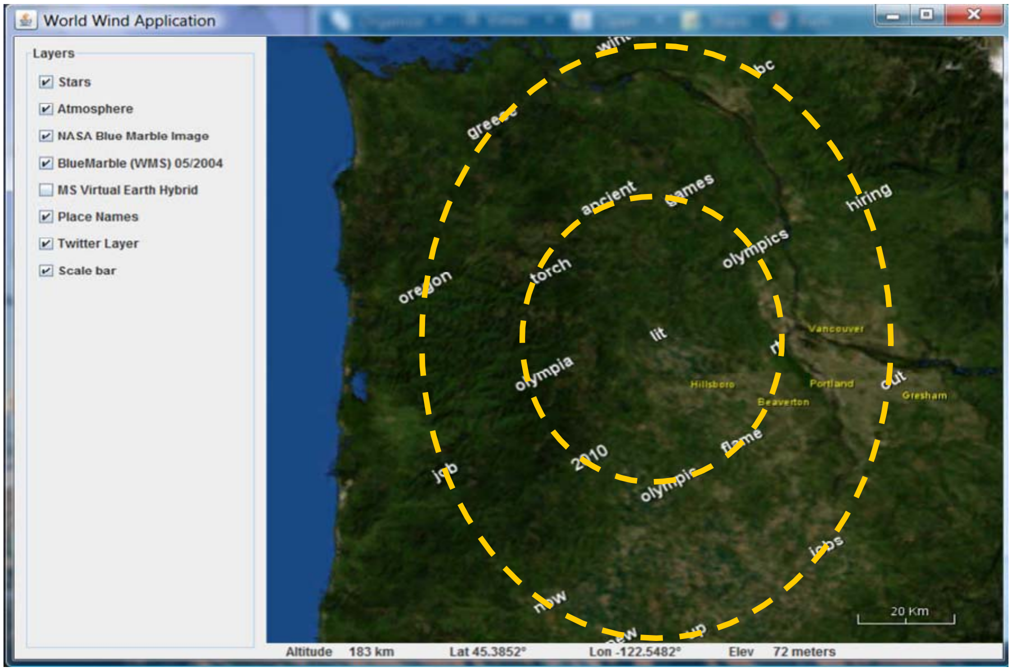


Figure 21: Example of a Word Cloud over Vancouver

- **Twitter Discourse Trends**

The second approach to showing Twitter data was to create a 3D visualization of the trends of the most important topics on the social networking site. The visualization is linked to live Twitter data, and it was created using OpenGL. We also explored adding the view in World Wind as shown below. While the view is presented in World Wind, it is not geo-spatially referenced, and it only provides a way to keep track of the Twitter trends while exploring GIS features.

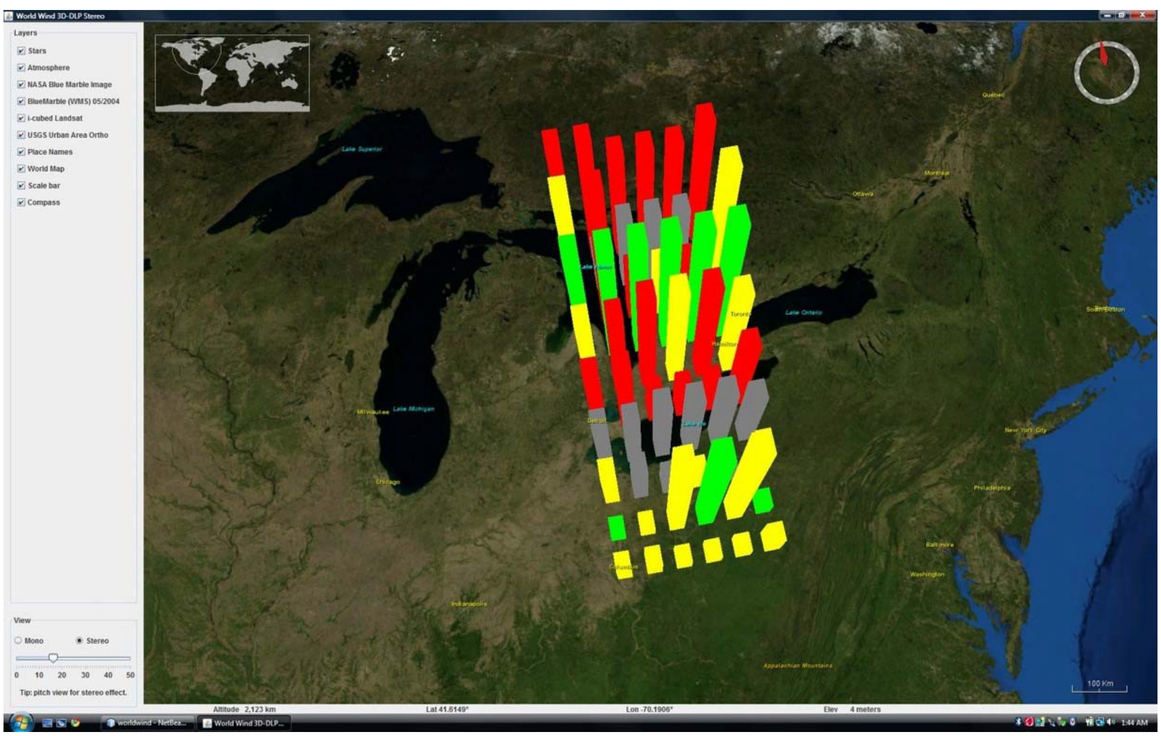


Figure 22: Trend Visualization of Twitter Discourse

3.1.3. 3D Models of Building in NASA World Wind

Summer at Tec^Edge also looked at introducing actual models of buildings into World Wind. In collaboration with Edgardo Molina and Rhonda Vickery, Track 2 also worked on developing a prototype for incorporating Collada files into the GIS environment. We integrated JCollada, an open source library for loading Collada (Digital Asset Exchange (DAE)) files, into World Wind. Java-based Collada (JCollada) uses the same rendering platform as World Wind, namely JOGL.

Collada models of Columbus, Ohio, buildings were downloaded from Google and placed at their coordinates. The prototype was not able to load larger models such as the ones provided by Woolpert and it also had problems with the materials (and colors) of some Google models. Other drawbacks of the approach include poor performance on a regular computer and manually not being able to recognize built-in Collada coordinates. These all seem to be related to the JCollada library being used.

Figure 23: 3D Model of the Ohio Statehouse in World Wind

3.2 Results and Discussion

3.2.1. Evaluation of Twitter on World Wind

Track Three is ready to conduct a full user study to better understand the advantages and drawbacks of our approach for visualizing non-textual information in a GIS environment. We are in the process of obtaining approval from the Wright-Patterson Air Force Base (WPAFB) Institutional Review Board (IRB) to start this experiment, which is the only step missing. We were already granted UALR IRB approval.

The study measures the effectiveness of Twitter on World Wind against a plain table view of Twitter, focusing on the relationships between Tweets and actual news. Activities completed for the study include:

- Designed a Java-based interface to record surveyor responses to each portion of the assessment across both platforms (Table View and Map View). Included sorting options on the Table View platform by City, State, and Twitter keywords; all these techniques are intended to make the Table View as powerful as possible, and thus an appropriate "competitor" to Twitter map layer.
- Archived World Wind place names and locations for cities and regions.
- Collected headlines from various news sources throughout the U.S. to test against Twitter keywords. Compiled Twitter keyword layers. Produced two alternate repositories of Tweets (ALT1 & ALT2).
- Developed specialized software to obtain Twitter data.
- Created a program to remove automated Tweets (e.g., Airport weather data).
- Linked the code for each platform (Table and Map Views) to the assessment interface.
- Prepared tutorial packets for survey participants.
- Designed notices to advertise the research study amongst the student body.

In preparation for the study, the researchers conducted an internal pilot study to better understand issues with World Wind, Twitter, and the user experiment. Some of our findings prompted us to look into developing advanced techniques for showing more precise distribution of Twitter data. While interaction with Twitter on World Wind can happen through "normal" techniques (that is, by panning and zooming) advanced techniques can pin-point the origin of any keyword and show exactly on the map other places into which that keyword occurs.

Initially, the technique was developed on the CAVE version of World Wind because the user can simply touch or point to the Twitter words of interest. We are in the final stages of implementing the technique in the desktop version of keyword, in which a simple mouse-over reveals distribution of keywords.

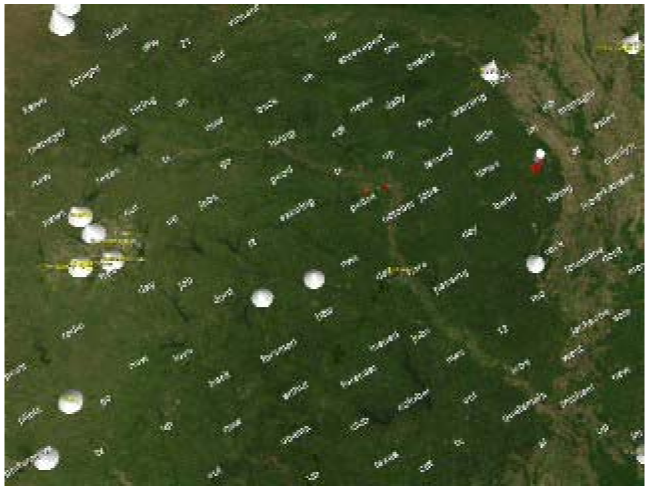


Figure 24: Advanced Interaction Example

3.2.2. Information Quality (IQ) Study of Twitter on World Wind

We present a case study to demonstrate the IQ aspects of integrating Twitter into World Wind in order to create a more enriched situational awareness for the user. More specifically, real time information from Twitter is logically and spatially displaced on a map so as to enable users to dynamically update their knowledge of a geographical region without having to manually create queries to extract information from Twitter, sift through tweets, or acquire new skills for utilizing the Twitter-World Wind system.

We analyzed the results of our automatically generated Twitter queries for keywords (words with highest frequency), sentiment (mood expressed in the tweets e.g. panic, sadness and happiness), and trend (rate of keywords and sentiment occurrence over time relative to geographical locations).

Part of our motivation for using this case study is to examine the feasibility and possible shortfalls of using Twitter as a source of real-time information in the decision making process of first responders. In this paper we focus on the IQ aspects of the study. We described the IQ issues using some of the data quality dimensions enumerated by Strong et al [9] namely believability, accuracy, timeliness, ease of use, interpretability and accessibility.

Timeliness**:** In today's information technology driven society, timely and effective response to emergencies and natural disaster is critical to first responders, not only because of the lives and livelihood that have to be saved, but because it also helps to promote and maintain good public relation and sometimes to preserve employment. An example was response by the Federal Emergency Management Agency (FEMA) to Hurricane Katrina in 2005.

We examined the lag time between the occurrence of an event and when an indication of it appears in our keyword analysis. Suppose that FEMA had been able to better integrate information from the news media and bloggers (most of the citizens learned about FEMA's inefficiencies though these media) with their other available information and displayed them spatially in a logical manner, it might have been possible to more effectively monitor and address the situations which contributed to the public backlash. Furthermore, the additional information might have saved more lives and livelihood.

We identified two factors that contribute to the timely appearance of an event's keyword(s) on the map. These factors are the number of tweets regarding the event and the size of the geographical region affected/interested in the event. Arguably, an event is tweeted about within seconds of its occurrence therefore it is plausible to assume that the knowledge can be transferred and viewed immediately in the GIS environment.

However, in order for keywords that signal an event to show up on the map, not only must it be tweeted about but the number of tweets about it must be significant enough to give its signal/buzz words a relatively high frequency. Therefore, important events will naturally surface to the map and overcome day to day tweets and spam (See Figure 5). The manner in which the keywords are displayed also provides a sense of the geographical distribution of the event. Users can perceive whether an event is generated over a large geographical area or whether the event is concentrated into a single spot by flying down into more narrow areas of the map as shown in Figure 6.

Therefore given an event like a hurricane, current trends indicate that within minutes the keywords about it would appear on the map. The Los Angeles Fire Department used Twitter to both inform and obtain from the public incidents of fire outbreaks [7] [8].

The new information product creates a more effective situational awareness by providing timely information that could help in decision making. We did not examine here the timeliness of other information such as place names and topography used in the GIS system, but that is likely to lag the timeliness of Twitter data because changes in place names or seasonal changes in landscape require some time to propagate, if ever, to the GIS databases.

Believability: We examine under the believability dimension the perceived integrity of the information from Twitter. Because information can be broadcasted by anyone on Twitter due to the very little restrictions available, there is ample opportunity for misinformation to be perpetuated. For this reason and because of the potentially critical decisions that would be based on our information product, believability is a very critical issue. To put it briefly, if users do not believe the information presented by the system, they would not rely on it for decision making, hence their decision making process is likely to remain the same or more complicated due to possible extra complexity introduced by Twitter.

NASA and US Geological Survey (USGS) GIS data available on World Wind have a high level of integrity, hence they are believable. The integration of Twitter into the GIS environment effectively creates an information product with both "high" and "low" believability. This integration can be used as an opportunity to increase the believability of the information derived from Twitter in certain cases. Essentially, by putting Twitter information in the context of location context its integrity can be more easily discernible and checked. For example, if a collection of tweets indicates the occurrence of flooding in the Mojave Desert, one may use this reason to easily conclude that such an event is not likely. Furthermore, if multiple contradictory events occur in the same location or neighboring locations, the user is better positioned to determine the believability of the information.

Generally, the more "reliable" information available to disambiguate the "less reliable" information, the easier it is to determine the believability of the "less reliable" information.

Accuracy**:** Under the accuracy dimension we analyzed how correctly the analyses of the tweets were displayed geo-spatially. The three factors we identified as affecting the preciseness of the geographical placement of keywords relative to the associated location are the focus altitude, proximity of places, and the number of words displayed.

The focus altitude is the distance above the sea level between the user's view and the earth's surface. Because altitude is a determining factor in the number and size of tiles used by World Wind to display surfaces, geo-referencing keywords is consequently affected. In effect, the same screen space is used to display information regardless of whether the view presents the whole US or just a single city. In the US view, the placement of Twitter keywords is less precise than in the view of the city because we take into account not only the ideal position of a keyword, but also its potential readability. As such, keywords may need to be moved around to provide enough inter-word spacing.

At very low altitudes, the tiles are small enough such that individual town/city on the PlaceName layers can be displayed on individual tiles. When places have proximity to one another, their displayed keywords sometimes overlap creating the possibility for wrong associations or confusions. The second and third factors are somewhat related because the overlap in keywords for different places is dependent on the number of keywords displayed for each place.

Ease of Use**:** Since many first responders' (e.g. military, red cross) command and control centers already use location based information in their operations, integrating new information in the context of location relatively helps in its assimilation. Furthermore, serving up Twitter information on an existing and familiar platform to the users concerned is preferable to having them master the intricacies of new applications. This can be very important when the time between learning to use new applications and responding to is very small.

In addition to presenting the new information on an existing platform (GIS), the ease of use of the information was also examined from the perspective of the amount of additional user activities/efforts required to operate the new system. It is therefore, desirable that the complexity of the system from this point is not increased significantly. Although the more technical details are not published here, the user is not required to do more than the usual pan, zoom and hovering needed to maneuver on World Wind. Twitter query is dynamically generated and the tweets are automatically analyzed and geo-referenced on the map.

The exploration of Twitter keywords in the GIS context is in fact completely free for the user, and no new skills are required. Flying through the Earth results in automatic filtering and aggregation or drill-down of the Twitter data.

Interpretability**:** This IQ dimension is improved for Twitter data because the user can at-a-glance see the overall structure of the keywords, can correlate tweets with GIS, and can discover new relationships between sets of tweets.

The display of the keywords shows information extracted from hundreds of tweets over an easy to interpret geographical milieu. The task of reading each tweet individually and of understanding the overall structure requires significantly more resources from a user.

Furthermore, even alternate forms of displaying the extracted keywords, such as tables, would still be harder to interpret and navigate (for example drill-down or increase the level of aggregation) than a map.

Another interpretability boost stems from displaying Twitter data in the context of a map. It is not unusual to have incomplete information in a tweet, sometimes by omission and sometimes because that information is self-evident to the sender. For example, there may be tweets that talk about an accident on the interstate south of XYZ. Without a map, it may be impossible to determine which interstate has the accident, but GIS data can simply disambiguate the highway.

Finally, placing keywords next to each other uncovers relationships between tweets that would be difficult to observe otherwise. Users can see events occurring in adjacent cities, states, and even countries. The spatial placement provides links between keywords that are not intrinsically written in tweets. Users can, for example, compare and contrast events taking place in Central Arkansas with events in the Fayetteville area.

3.3 Conclusions

Track Two successfully created new techniques for improving the IQ of the Layered Sensors Domain. Our techniques are modular in nature and can be combined with each other or with other World Wind layers.

The presentation of data in a CAVE can let the user feel immersed in it, and better understand 3D relationships. The demonstration at Tec^Edge required only a few hours to set up a portable immersive system. The software included both the CAVE port and Twitter on World Wind.

The integration of Twitter data with GIS data promises to increase the quality of both data sources. Twitter has very good timeliness, but may suffer from accuracy, noise, and believability, while GIS data is largely correct, but slightly outdated. Combining the two sources provides an overall increase in quality because one can draw on the other's strengths.

3.4 Introduction Year Two Work

The second year of work on IQ Tools for Persistent Surveillance Data Sets extended research done during the first year. Track Two research was extended to include data collection, cleaning, and archival, to integrate CAVE and iPad, and to analyze text and develop desktop visualization.

3.5 Objectives

The Track Two objective was to continue improving the IQ of the layered sensor domain by integrating, managing, and presenting non-spatial, textual data. Three quality dimensions, accessibility, relevance, and representation where improved in that:

- We provided a potential analysts access to large amounts of textual information that could be easily analyzed in the context of spatial layered domain (tasks a-d);
- We developed methods of processing and discovering relationships in the information, while allowing the analysts to navigate through the relevant information (tasks e, i-m); and
- We created new presentation methods for the information to alleviate the high demand placed on an analyst (tasks f-j).

A large thrust of the Track Two work focused on developing and investigating the quality of a 3D visualization, hence named Buzz Vizz, in a layered sensor domain. The principle behind the Buzz Vizz research was to analyze whether the proposed visualization could aid user's to improve his/her situational awareness during an event. By combining non-spatial data with GIS information our research attempts to analyze if there is an improvement in the confidence and efficiency of information based on this data fusion technique.

Besides BuzzVizz, Track Two developed TreeMap based textual layers, and began designing a new visualization (Chat Magnet) that analyzes the semi-structured transactions of social networks, particularly chat stream data, to improve our understanding of relationships and a user's state of being within electronic conversations, e.g., online forums, micro-blogs, iRC chats, blogs, etc.

Track Two research included the following tasks:

- Designing software for archiving Twitter and news feeds as well as supporting dynamic queries/analysis
- Implementing the capability to archive live Twitter data
- Creating a database capable of holding non-spatial data sources
- Extending archive capability to allow parsing and archiving of news sources
- Designing and implementing a layer capable of showing temporal aspects of non-spatial data
- Developing interaction techniques and assoc. visualizations that
    - ➢ allow for the control of detail presented through a layer
    - ➢ control the temporal aspect of the visualization layer
- Porting the 3D visualization of Buzz Vizz for use in the CAVE
- Porting the 3D visualization of BuzzVizz to the Apple iPad based on a remote portal of Buzz Vizz (this task pre-empted incorporating the height dimension of Buzz Vizz)
- Collecting experimental data to test the effects of Information Quality (IQ) for the Twitter layer of Buzz Vizz
- Analysis of the collected data
- Documenting the analysis process and interaction techniques for publication
- Developing software for the analysis of computer mediated communication (CMC)
- Researching datasets and functions for use in social network analysis (SNA)

3.6 Methods, Assumptions, and Procedures

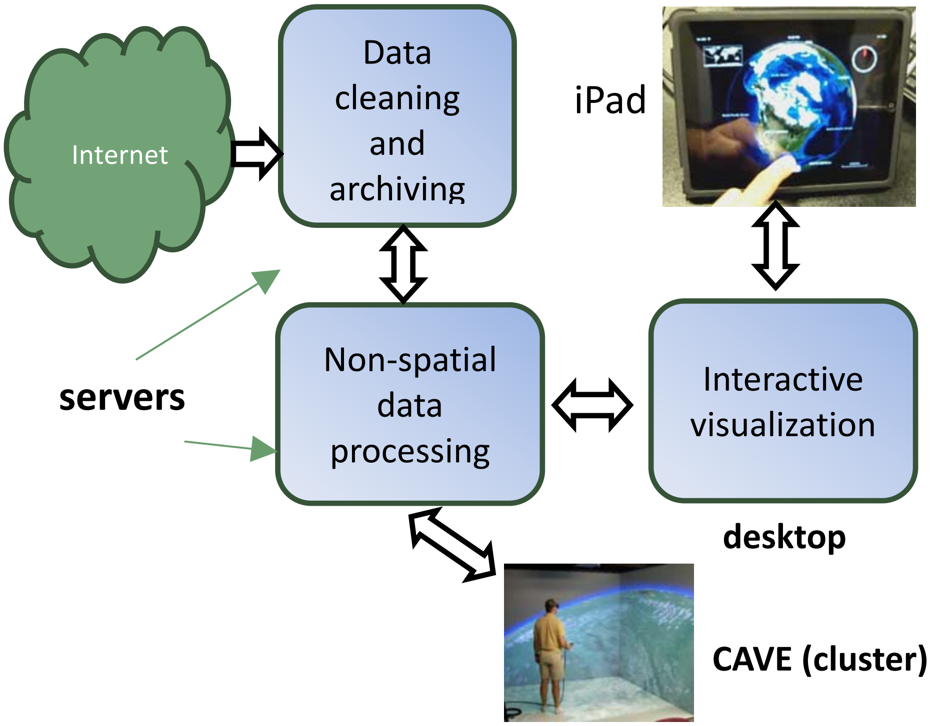


Figure 25: The Distributed Software Architecture of our Approach

Track Two research determined based on the findings of Year One, that, in order to improve Information quality in the three dimensions mentioned above (accessibility, relevance, and representation), two main problems needed to be solved. First, computational power is needed to deal with the large amount of textual information that was streaming from the web, which led us to modify the architecture of our approach to what is presented in Figure 25. Second, the reaction of the human users of our system needed to be explored in order to understand the how various features help users and how to fine tune them in the future. An experiment involving users of our system was conducted and analyzed.

The rest of this section explores : (I) Data cleaning and archiving; (II) Text analysis and desktop visualization; (III) CAVE and iPad visualization; (IV) Empirical study to research issues associated with interactive four dimensional (4D) information; and (V) Document the analysis process and interaction techniques.

The early stages of a new research subject was also established during Year Two, which looks to analyze the effectiveness of determining relationships between users within social media networks, e.g. iRC, chat forums, blogs, etc. via Social Network Analysis (SNA).

3.6.1. Data Cleaning and Archiving

This portion of Year Two work includes the development of software capable of collecting textual data from the Internet and the creation and management of a large (hundreds of millions of items) database archiving data since October 2010. Twitter information is by far the largest data store and it was collected via two main mechanisms, fire hose streaming and active querying for carefully chosen places and terms. Besides Twitter, the data we collected includes news feeds and chat data.

To stream the tweets, a Java program capable of connecting to twitter API and extracting the data was developed. The extracted data is composed of a row data that needs to be cleansed and structured, thus improving its quality. A relational database capable of storing the streamed data was created; the Twitter database comprises three tables Place, Status, User as shown in Figures 26, 27, and 28, respectively. The java program and the Tweeter database (MySQL database) were installed on a server and running 24 hours a day resulting on an upload of more than 250 million tweets at a rate of 760,000 tweets per day producing an accumulation of 120 gigabytes (GB) of data.

A web crawler was developed using java programming language to extract chat logs and store it in a file. The file then was uploaded to the database (chatlogarch that holds one table chatlog shown in Figure 30) where multiple processes of data cleansing were run using SQL query and Java Programming Language.

Another Java program was developed to fetch for Really Simple Syndication (RSS) feeds from the web and stores them in MySQL database (RSS Table, Figure 31).

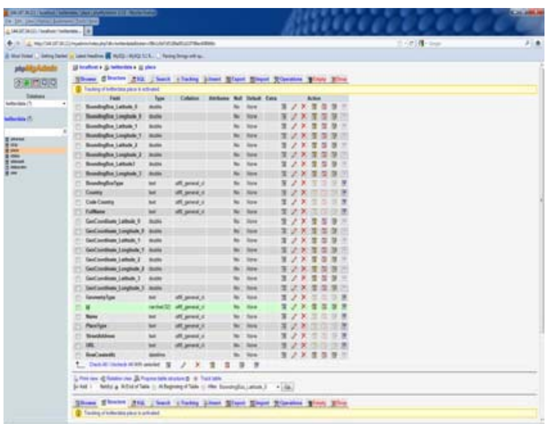


Figure 26: Place Table

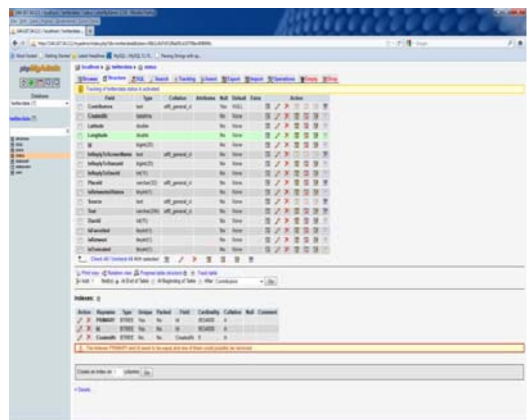


Figure 27: Status Table

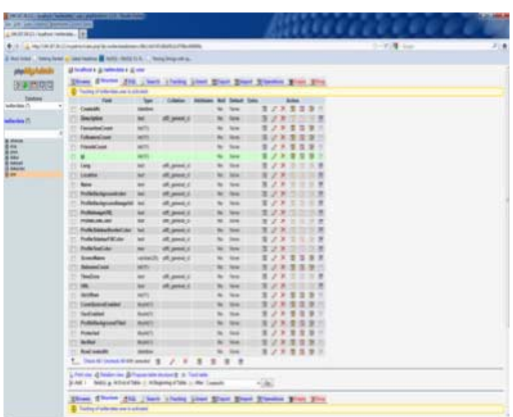


Figure 28: User Table

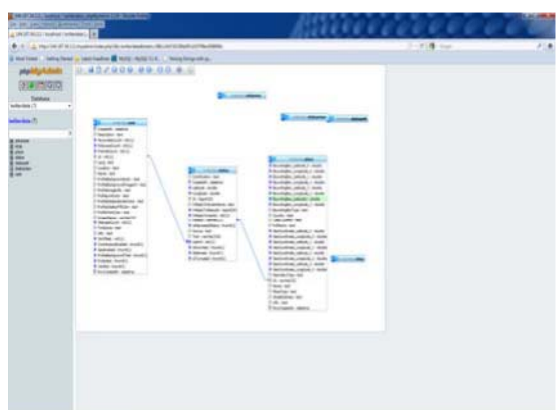


Figure 29: Twitter Relational Database

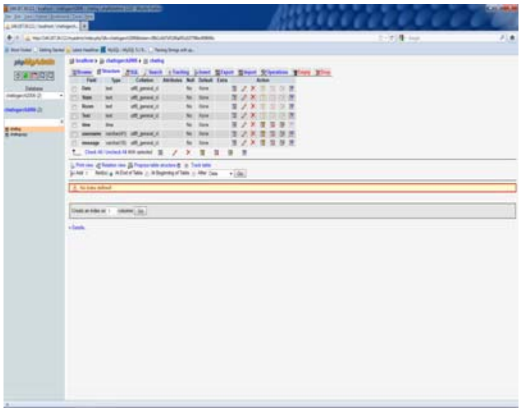


Figure 30: Chatlog Table

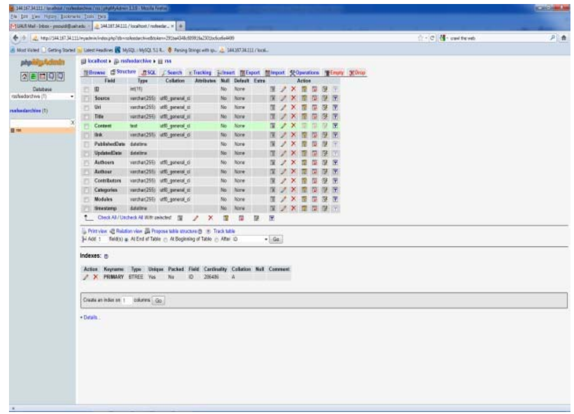


Figure 31: RSS Table

3.6.2. Text Analysis and Desktop Visualization

The work performed here involved the modification of the architecture of the tool to allow it to take advantage of multiple machines as depicted in Figure 25, the introduction of time dimension and controls to allow users to see how the discourse on the web changes over time, and the additions of additional features that allow the analysts to filter to the relevant information and to see who, not only what, participates in the discourse. We also developed another layer representation that uses a TreeMap instead of a 3D map as the backdrop for tweets information. The TreeMap can display hierarchical information and organizes Twitter keywords according to this information.

The text analysis portion of the system was re-designed and implemented in such a way to allow it to run independently of the visualization and data archive. Network communication protocols were integrated into the software to allow the three main pieces to run on multiple machines. Text analysis can work seamlessly with either desktop visualization or CAVE immersive navigation.

The analysis was improved with additional association algorithms (data mining), and with the capability to extract and process Twitter users, not only Twitter keywords. A filtering mechanism based on low-level tweet content was added in order to allow the visualization to filter out tweets that are not relevant to the user's interest. Finally, the analysis was made aware of the temporal aspect, which can be extracted from the archived data over time.

The visualization was added to temporal controls as shown in the lower left corner of Figure 32. The analyst can make Twitter data play over time. To better see the temporal relationships between keywords, they are given different sizes, with the most current one being the largest. As such, as a theme tapers off over time and it becomes smaller and smaller until it disappears. We found out that about five sizes can be distinguished by a viewer and because each time step is one hour, the viewer can at any point see the last five hours of tweets over any geographical area.

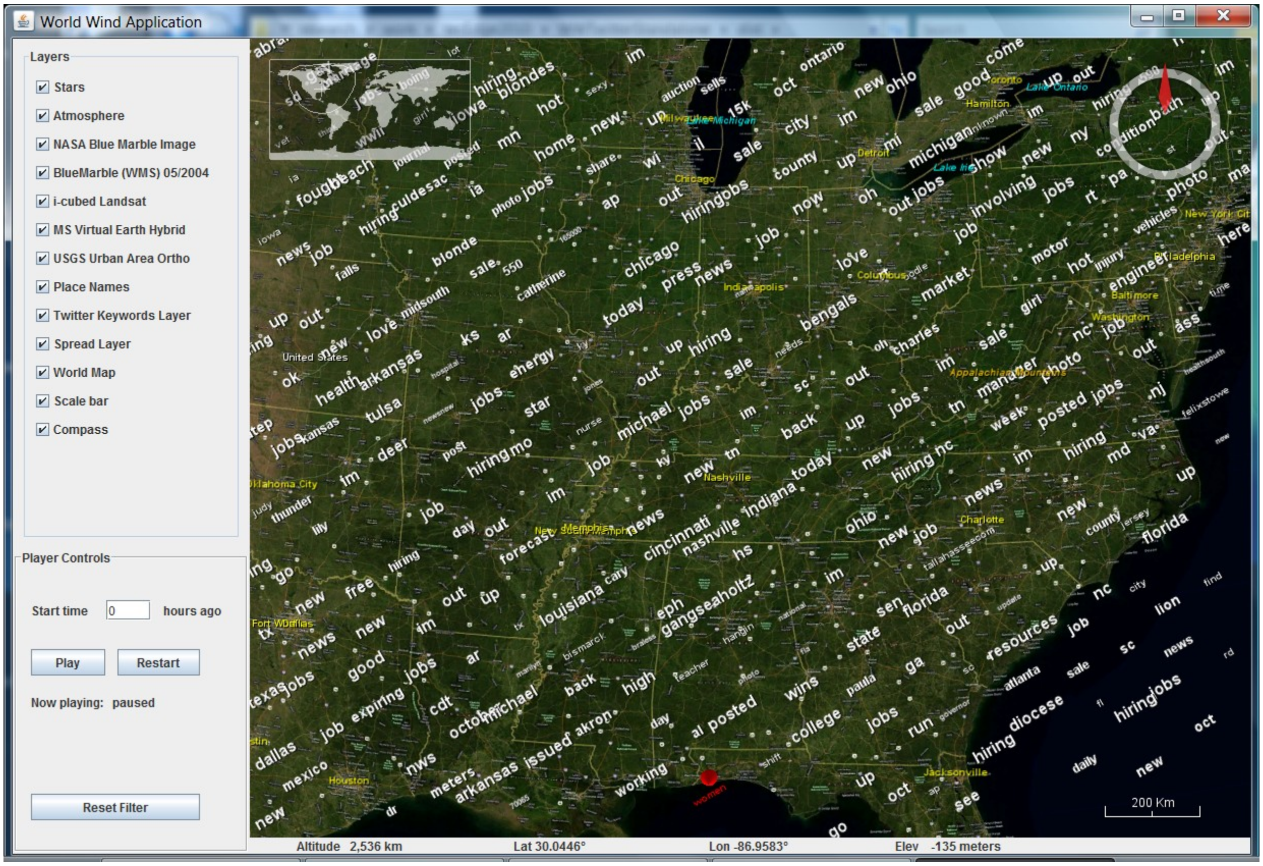


Figure 32: Desktop BuzzVizz with Temporal Controls Shown in the Lower Left Corner
(Twitter Keywords of Different Sizes Show How Much Time has Elapsed Since that Word was Last Tweeted in a Particular Area. The mouse is hovering over the red "women" keyword, and the spread) layer shows where that word was tweeted from red cone)

BuzzVizz was also enhanced with the following two features in order to improve the relevance and accessibility dimensions of IQ:

- Clicking on a twitter keyword, requests the analysis be limited to only those tweets that contain that feature. Users can click on multiple keywords, one after the other, and only those tweets containing the selected keywords are included in the analysis. The filter words are displayed with a different color. Clicking on a filter word again, removes it from the filter. A "Reset Filter" button is added to allow all filters to be removed at once.
- A User Layer is added to BuzzVizz. This shows the most frequent Twitter users over a geographical area. This layer has (as the regular keyword layer) a Spread Layer associated to it, which allows the user to mouse over a user name and see instantly the cities associated to that user as a set of red cones on the map.

The final part of work included the implementation of a Twitter layer on top of a tree visualization (Figure 33). This would allow an analyst to look over the structure of an organization (usually hierarchical data) and determine what has been Tweeted about different areas of that organization. Clicking on a branch of the tree would result in only that branch being shown (regular TreeMap behavior). As such, the users can drill-down and zoom out as needed. We noticed that it became difficult sometimes to distinguish the words due to some color issues.

Figure 33: Twitter Layer on a TreeMap

3.6.3. CAVE and iPad Visualization

The main activity for the CAVE visualization was to take advantage of the new architecture, and refactor the code in such a way to reduce the differences between the code used for the CAVE and that used for the desktop. The main problem in Year One's software was that a custom interaction component was used for the CAVE, which required a lot of duplication between the desktop and CAVE versions. Features added in the desktop will have to be custom re-coded for the CAVE.

The current architecture and implementation uses an abstraction for the interaction classes, and makes all features coded for one type of display easily transferable to the other. The CAVE now has all the features of the desktop, and any future feature in one can be ported effortlessly into the other as shown in Figure 34.

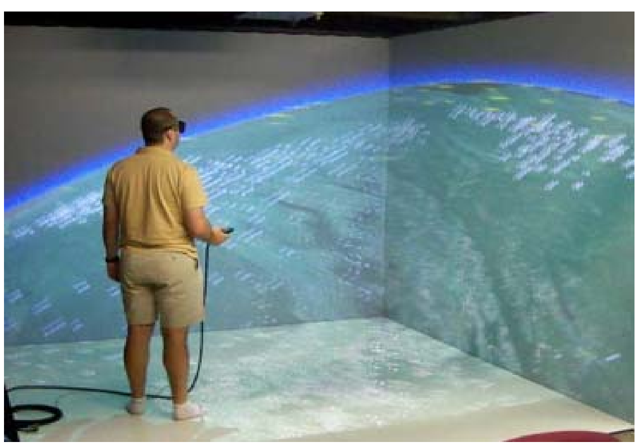


Figure 34: One of the Investigators inside the CAVE Running BuzzViz

The implementation of the BuzzVizz to Apple's iPad was challenging because the operating system of the iPad (iOS) does not support Java, which is the language behind BuzzVizz. Our tool is based on hundreds of thousands of lines of code from NASA World Wind, and a total re-coding was beyond our men-power and time frame.

The solution employed by Track Two consisted on creating a remote display of a BuzzVizz instance running on a regular desktop. To achieve this, a capture module was added to BuzzVizz, and a remote display and interaction application was installed on the iPad. The user would in fact see and interact with the version of BuzzVizz running on the desktop. The communication of the image stream from the desktop and of the user interaction from the iPad takes place over a wireless network. We employed a number of techniques to ensure a smooth and fast remote interface. Figure 35 shows a series of snapshots from a user interacting with the iPad.

Figure 35: BuzzVizz on the iPad

3.6.4. Empirical Study to Research Issues Associated with Interactive 4D Information

An empirical user study was designed to test the team's hypothesis that the BuzzVizz visualization was superior to just looking at Twitter via a table showing tweeted words per city and state (table platform) on several IQ dimensions (accuracy, believability, etc). The study design took into consideration ***place***, ***headline***, and ***platform*** as independent variables. ***Time***, ***keyword***, ***confidence***, ***user headlines***, and ***user selected place*** were equated as dependent variables. The test software randomly chose cities and regions within North America for participants to focus on for answering questions within each platform. The five-part survey contained questions relevant to determining:

- User background
- Ability to summarize headlines from a given region or city
- Ability to choose locations of headlines based on real or synthetic news
- Ability to list keywords common to cities in specific regions
- Any improvements to the visualization

In terms of background details relevant to the study, a majority of participants (82%) indicated they received news via the Internet at a rate of 68% per week. The background analysis further revealed that only a small percentage of participants follow news from social media blogs like Twitter (17%). This could be in large part due to Twitters reputation as a viable source of information (55%) compared to such scientific institutions like NASA (95%).

The results of all participant responses (22) to study questions related to genuine news data were verified alongside independent news sources, i.e. LexisNexis and analyzed against select IQ dimensions for ease of summarizing headlines, confidence in performance, ease of navigation, ease of finding cities and regions, platform and time performance. Figure 36 illustrates the adjusted comparison of the visualization tools against the objective evaluation of headlines.

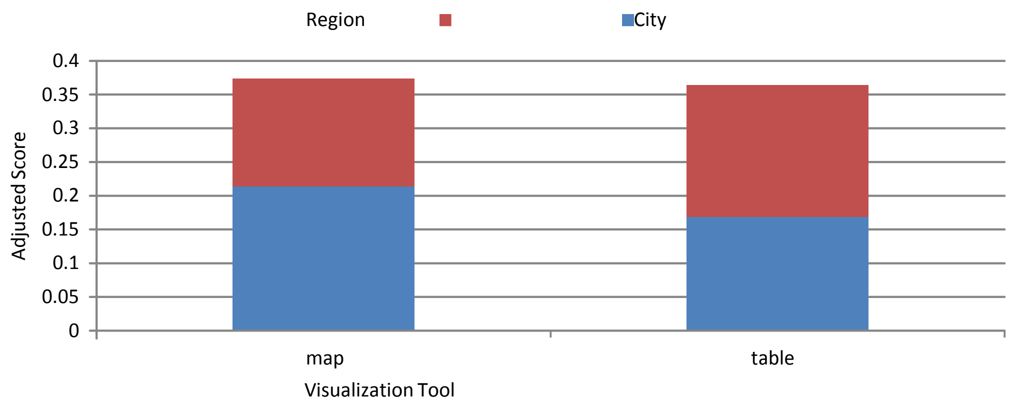


Figure 36: Adjusted Score Comparison of Each Visualization Tool

The accuracy of participant responses to portions of the study were adjusted based on the **place** variable (selected by the user or randomly provided by the program). Our research also analyzed participant's accuracy of answering questions related to listing keywords tasks by region, which resulted in a 50% performance improvement when using the map platform. Figure 37 illustrates these findings based on the adjusted score average and time average.

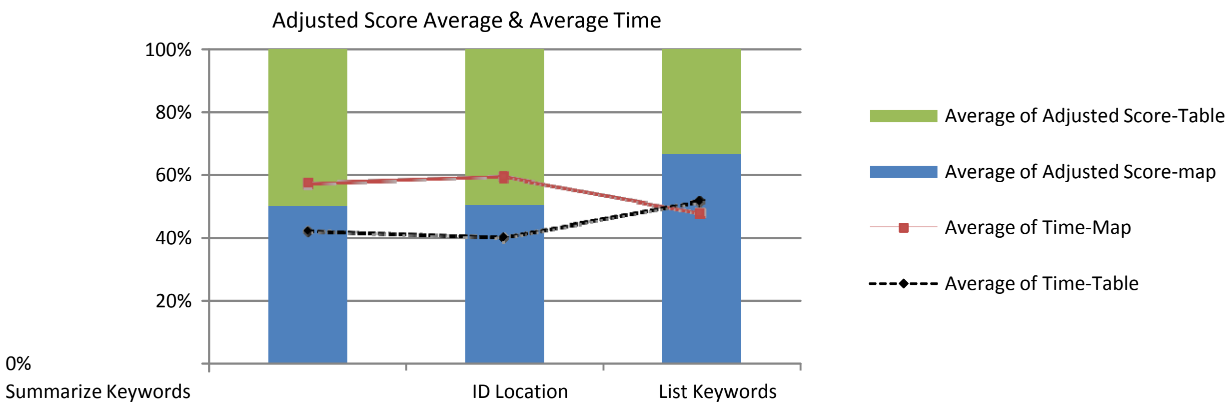


Figure 37: Summary of the Accuracy of Region-Related Tasks by Category

In addition to the higher accuracy achieved on the map for the keywords listing task, there is also better time performance for the task on the same visualization tool (8%).

Our next analysis compared the reputation of Twitter information (a component of the information product) expressed by the participants in Part One of the survey to their belief in the summarized headlines generated from the final information product (visualization). Initially, participants had a 55% believability in Twitter data, but it was later determined that participants had a 75% believability in information based on the average believability indicated by summarized headlines.

Figure 38 illustrates that participants' previous doubts about the believability of Twitter information did not affect their believability of information on BuzzVizz. The believability responses for the map and table are almost identical suggesting that the blending of the information sources rather than the visualization is responsible for participants' relatively high confidence in the final information product.

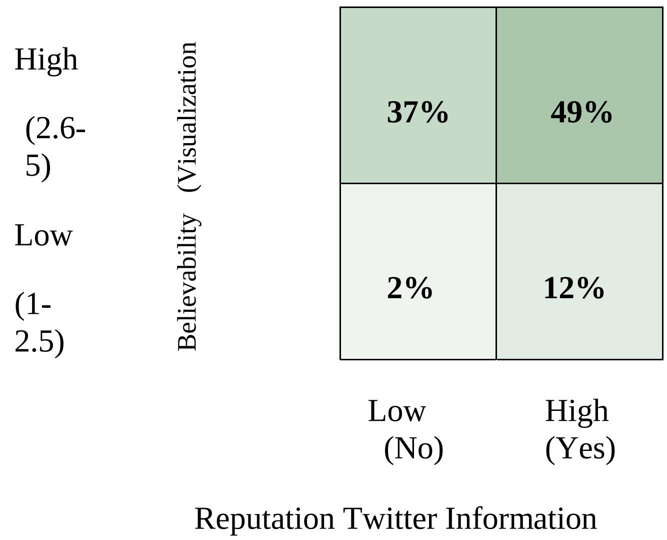


Figure 38: Subjective Assessment of the Information Product and its Component

Tables 4 and 5 below summarize broad comparison of the performance of the map and table presentations based on tasks categories (Part 1, 2, and 3) as well as region versus city. Dark gray encodes similar performance; light gray represents table, and white shows BuzzVizz/map.

Table 4: Broad Comparison between the Table and Map Visualizations

	Average of Adjusted Score	Average of Time	Average of Believability Scale
Region	Table	Table	Table
City	Map	Table	Map
Summarize	Same	Table	Map
Find Place	Map	Table	
List Keyword s	Map	Map	

Table 5: Counts of Tasks Categories of Better Performance

Visualization	Average of Adjusted Score	Average of Time	Average of Believability Scale	Σ Performance
Map	3.5	1	2	6.5
Table	1.5	4	1	6.5

The map and the table visualizations appear to have the overall equal efficiencies. The map leads on accuracy and believability while the table holds a substantial lead on the time performance. However, if the learning curve is taken in to consideration, the time performance advantage enjoyed by the table is expected to be diminished or eliminated over time.

Table 6: Distribution of Participants' Preference on Visualization

	Map View	Table View
Ease of Summarizing Headlines	11	11
Confident in Responses	11	11
Ease of Navigation	13	9
Ease of Locating Cities/States	8	14
Task Platform Preference	14	8

The equal number of participants who indicated confidence in their responses on both visualization platforms correlates to the almost identical believability response in the headline summarization task. The distributions of participants regarding the ease of navigation and the ease to locate cities/states are conflicting. Further analysis indicated that six and seven participants found it easier to navigate and locate cities/states on the map and table respectively while two participants found it easier to navigate on the map but more difficult to locate cities/states on it. The remaining seven participants found it easier to navigate on the table but more difficult to locate cities/states on it. 64% of the participants preferred the map to the table for the given tasks.

3.6.5. Document the Analysis Process and Interaction Techniques

Track Two successfully submitted an article entitled "A Study of the Quality of a Visualization that Employs a Geospatial Milieu to Convey Twitter Data" for review to IEEE Computer Graphics and Applications (CG&A) 2012-1 (Visualization Applications and Design Studies for IEEE CG&A January/February 2012 issue). The team is planning on submitting more articles to upcoming conferences based current research in social network analysis and its corresponding visualization Magnet Chat.

3.7 Chat Magnet

Track Two took initiative to develop a visualization technique that would allow a chat user to determine what topics have been discussed in the past. Our approach uses a set of magnets that "extract" information based on their type. Magnets can pull keywords or user names from the chat if those are related to the type of magnet. For example, a "Problem" magnet can interact with the chat messages that refer to a problem, while a "Supply" magnet could interact with those messages that promise something to be delivered. By using these magnets, the analyst can determine who had a problem and who promised to solve it.

Chat Magnet has two areas: a chat area and a magnet area. They are both zoomable user interfaces and as such they can be smoothly zoomed in and out and panned. The chat area can only be panned vertically which corresponds to moving back or forward in time. A snapshot of Chat Magnet is shown below.

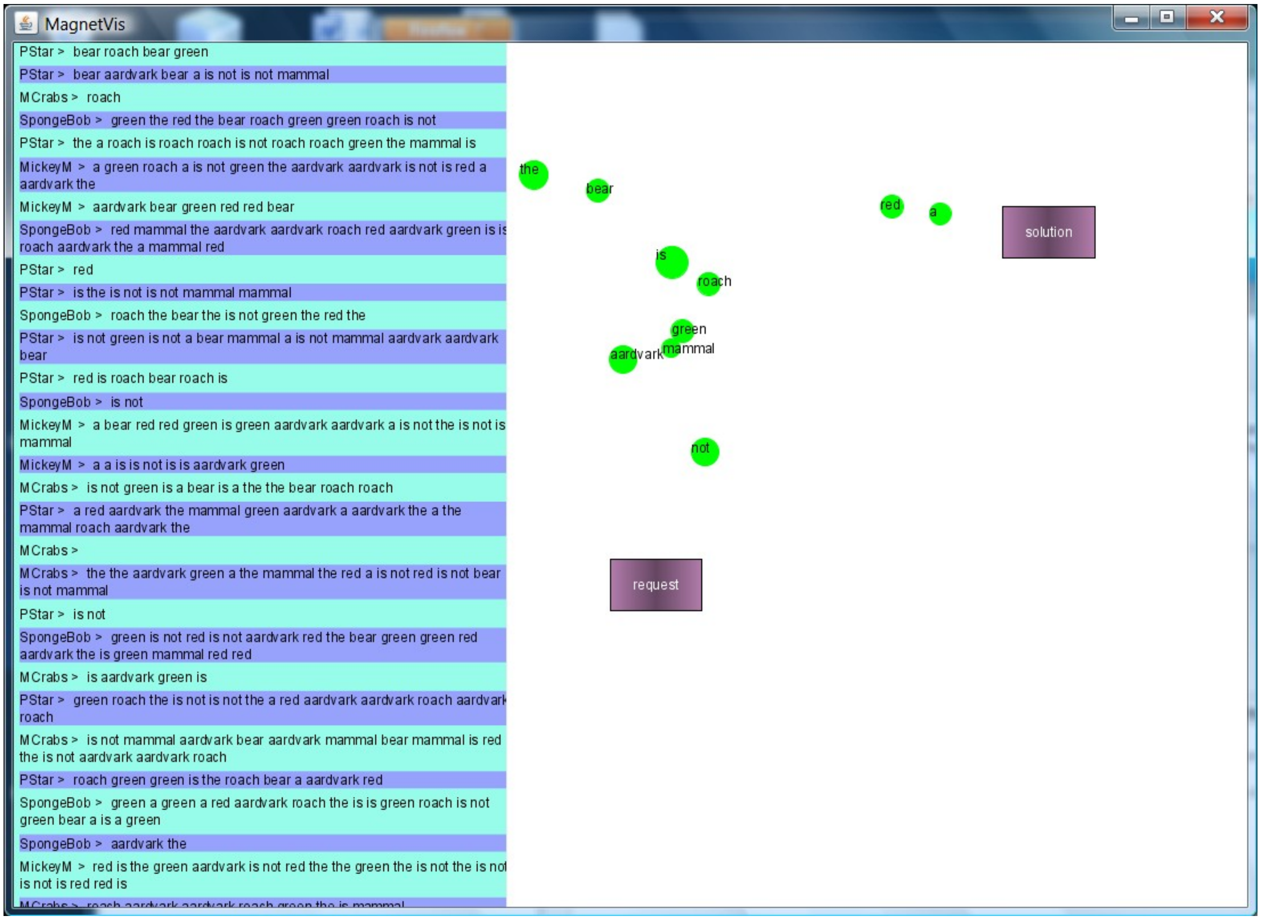


Figure 39: Chat Magnet Visualization

(Chat Area is on the left, magnet area on the right is showing two magnets)

We examined a set of visualization tools and related ideas are introduced for extracting and analyzing social network features present within persistent, multi-speaker, multi-topic, quasi-synchronous computer-mediated communication systems enacted over the internet, i.e., internet chatrooms.

Our approaches to analyzing the social network include:

Keyword Similarity: Messages were assigned a similarity score based on the number of keywords they share messages. Shared keywords are considered as informative of both the substance of the conversation and like likelihood of user interaction.

Temporal proximity: Known temporal pattern of human communication was used to estimate the probability that later messages by other users is a response to an earlier one by a particular user.

Direct Addressing: This feature of digital text communications is used by speakers to enhance the clarity of who the message is intended for (and who it might not be intended for). It usually involves the mentioning of the intended recipient of the message.

A visualization idea being considered is similar to a covariance matrix (but using relative frequency of contacts instead of co-variances). As shown in the figure below, individual communicators and their connectedness to others can be quickly visualized without reading through several lines of messages. One disadvantage of this method, again, is that for very large numbers of communicators, these graphics would get complicated very quickly.

Jane	Mike	Bob	Eric	Mary	Shelly	Fred	Joe	Kate	Lindsey	Martha	Paul	George	
	100	27	32	0	23	28	23	23	24	0	0	21	Jane
100		30	27	32	24	48	24	51	51	25	23	22	Mike
27	30		26	0	0	21	0	0	0	0	21	0	Bob
32	27	26		0	21	0	0	0	0	0	0	0	Eric
0	32	0	0		0	26	0	0	23	0	0	0	Mary
23	24	0	21	0		22	0	21	22	0	0	0	Shelly
28	48	21	0	26	22		30	34	53	0	25	0	Fred
23	24	0	0	0	0	30		0	26	0	0	0	Joe
23	51	0	0	0	21	34	0		44	23	21	0	Kate
24	51	0	0	23	22	53	26	44		0	23	0	Lindsey
0	25	0	0	0	0	0	0	23	0		0	0	Martha
0	23	21	0	0	0	25	0	21	23	0		0	Paul
21	22	0	0	0	0	0	0	0	0	0	0		George

Figure 40: An Example Contact Matrix

(This displays who is talking to whom and with what relative frequency)

3.8 Conclusions

Track Two was successful in analyzing the empirical study, and establishing that BuzzVizz has a number of benefits to the user and it is the only one capable of showing geographical data. New features added to BuzzVizz are likely to make it more useful. Finally, the ability to use BuzzVizz on a number of platforms ranging from Apple iPad, to immersive CAVE environments makes our approach versatile.

We tried an alternative to BuzzVizz based on using hierarchical data instead of a map, but for that to be a viable option, new coloring techniques for Twitter keywords need to be developed. Hierarchical information is useful because it expresses data covering typical organizations such as corporations and governments.

4.0 INTRODUCTION TO TRACK THREE – VISUAL RENDERING AND DISPLAY OF TEXT & IQ METRICS – SMART ENVIRONMENT

Smart environments refer to buildings or locations equipped with a multitude of sensors and processing mechanisms for improved security, efficiency or functionality. Often, these sensors serve distinct purposes and their data may be processed separately by entirely separate systems. We argue that integrated processing of data available from multiple types of sensors can benefit a variety of decision making processes. For example, smart building sensors such as occupancy or temperature sensors used for lighting or heating efficiency can benefit the security system, or vice versa. Recent industry standards in sensor networks such as ZigBee make it possible to collect and aggregate data from multiple, heterogeneous sensors efficiently. However, integrated information processing with a diverse set of sensor data is still a challenge.

We provide an information processing scheme that offers data fusion for multiple sensors such as temperature sensors or motion detectors and visual sensors such as security cameras. The broader goal of multi-sensor data fusion in this context is to enhance security systems, improve energy efficiency by supporting the decision making process based on relevant and accurate information gathered from different sensors. In particular, we investigate a major data fusion technique, Bayesian network, and present a simulation tool for a "smart environment." In addition, we discuss the potential impact of data fusion on the processes of decision or detection, estimation, association, and uncertainty management.

One of the outcomes of data fusion is the improved information quality that assists various decision making processes in a "smart environment." Our focus here is the integration of sensors information into the real-time decision making process in a surveillance context. We use data fusion in a fashion where different types of information are collected from a heterogeneous set of visual and non-visual sensors. The process of integrating data from different sources requires designing an appropriate data fusion model that would take the sensor data, integrate them following a certain model, and transform it to a set of useful and relevant decisions. The anticipation is for the resulting decisions to be more accurate and efficient than those resulting from a single source. In a broader sense, we expect data fusion to lead to a virtual collaboration between the different collected information.

Towards this goal, we first investigate the usefulness of data fusion in a smart environment equipped with visual and non-visual sensors and design a convenient data fusion model. Then, we provide an overview of data fusion methods, present our data fusion algorithm and discuss our data fusion engine. This is followed by a description of our smart environment simulation tool which is used to test some of the hypotheses, visualize the environment with the sensors and their spatial relationships and to allow us to build some of the case scenarios which is discussed last. In the last section, we summarize our findings and conclusions with a set of ideas for ongoing work.

4.1 Methods, Assumptions, and Procedures

Data fusion is "the theory, techniques and tools which are used for combining sensor data, or data derived from sensory data, into a common representational format." Fusing data from different sources can improve the quality and the utility of information and help improve efficiency, security and functionality. The critical problem in multi-sensor data fusion is to determine the best procedure for combining information from different sensors in the system.

Most of the reported work in data fusion uses a statistical approach in order to describe different relationships between sensors taking into account the underlying uncertainties [4]. Edward Waltz and James Llinas summarize the methods to implement data fusion as follows: decision or detection, estimation, association, and uncertainty management theories. In decision or detection theory "measurements are compared with alternative hypotheses to decide which ones best describe the measurement." Basically, the decision theory assumes "the probability descriptions of the measurement values and prior knowledge to compute a probability value for each hypothesis." [2].

Fuzzy logic, neural networks, Bayesian, and Dempster-Shafer theories are the most commonly used methods in multi-sensor data fusion. However, our approach will focus on Bayesian model for integrated information processing using data from multiple, heterogeneous sensors. The main reasons for this election were the appropriateness of the input and output types in Bayesian model and its wide-spread use for similar problems in the literature. We plan to expand our work into the alternative fusion techniques as part of our ongoing research.

The basic principle of *Bayesian theory* is that all the unknowns are treated as random variables and that the knowledge of these quantities can be represented by a probability distribution. In addition, Bayesian methodology claims that the probability of a certain event represents the degree of belief that such an event will happen. The degree of belief is associated with a probability measure that can be updated by additional observed data. All the new observations are added to update the prior probability and therefore obtain a posterior probability distribution [3].

4.1.1. Bayesian Data Fusion

The Bayesian model integrates data, independently, from r correlated sensors' inputs in the following pattern:

$$\mathrm{p}(\mathrm{D}/\, X_1^1 X_1^2 \ldots X_1^r) = \frac{\prod_{j=1}^{r} p(D/X_1^j) * p(D/X_0^1 X_0^2 \ldots X_0^r)}{\prod_{j=1}^{r} p(D/X_0^j)} * \mathrm{K}$$

where K is the Bayesian normalization and is equivalent to $\frac{\prod_{j=1}^{r} \mathrm{p}(\mathrm{x}_1^{\mathrm{j}}/\mathrm{X}_0^{\mathrm{j}})}{\mathrm{p}(\mathrm{x}_1^1 \mathrm{x}_1^2 \ldots \mathrm{x}_1^{\mathrm{r}}/\mathrm{X}_0^1 \mathrm{X}_0^2 \ldots \mathrm{X}_0^{\mathrm{r}})}$ and $\mathrm{p}(\mathrm{D}/\, \mathrm{X}_1^1 \mathrm{X}_1^2 \ldots \mathrm{X}_1^{\mathrm{r}})$ is the probability of event D given $\mathrm{X}_1^1, \mathrm{X}_1^2, \ldots, \mathrm{X}_1^{\mathrm{r}}$.

$\mathrm{x}_1^{\mathrm{j}}$: Current measurement/observation from correlated sensors j where j = 1, 2, …,r.
$\mathrm{X}_0^{\mathrm{j}}$: Prior information or old data set from correlated sensors j where j = 1, 2, …,r.
$\mathrm{X}_1^{\mathrm{j}}$: Posterior information or new data set from correlated sensor j where j = 1, 2, …,r.
D : Event in question, i.e. one of the decisions labeled on Figure 41 below.

The fusion engine in this project is the model we use to integrate information from both visual and non-visual sensors. The engine we design receives inputs from both visual and non-visual sensors and provides a set of relevant decisions (outputs).

As the diagram in Figure 41 shows, s_1, s_2, s_3, … s_n are inputs from different non-visual sensors. These inputs first go through a correlation model (shown as raw data processing in Figure 41) that determinates the correlations among the sensors' inputs and transmits independent *m* outputs that are fed to the fusion engine as inputs. These outputs (fusion engine inputs) are labeled as

$x_1, x_2, x_3, \ldots, x_m$.

The fusion engine inputs $x_1, x_2, x_3 \ldots x_m$ can be matched to notations such as $X_1^1, X_1^2, X_1^3 \ldots, X_1^r$, which represent the posterior information, described in the algorithm section, from correlated sensors. However, this matching does not restrict matching x_1 to X_1^1, x_2 to $X_1^2 \ldots$etc as the data fusion model we use consider integrating posterior information from both non-visual and visual sensors. As it is explained below, data from visual sensors is pre-processed before it can be fed to the fusion engine. This pre-processing results in a convenient format of information to be passed to the fusion engine.

For visual sensors, we use optical and infrared cameras to record raw videos. The acquired videos are then processed to extract meta-data information to be used in the fusion algorithm described above. The processing of images from such visual sensors requires a preliminary processing where some intermediary image features such as moving objects and their boundaries are extracted for further processing [5]. The final extracted visual information forms metadata that can be fed to the designed fusion engine that integrates it with other sensor data from other heterogeneous sensors.

The extraction of visual information can be a real challenge because of “the lack of proper low-level algorithms for robust feature extraction” [7]. Here, we use a motion detection algorithm to extract relevant visual information about the moving objects in the recorded video. The algorithm chosen for this purpose is the implementation in OpenCV, which is an open-source computer vision library, originally developed by Intel. We have performed a few modifications at the input level that resulted in movement detection. The metadata in this context includes the kind of information such as the number of moving objects, the nature of movement, the type of the moving objects (human or animal), the actions performed by the moving objects, the area they occupy, and the time they stay in the room of question.

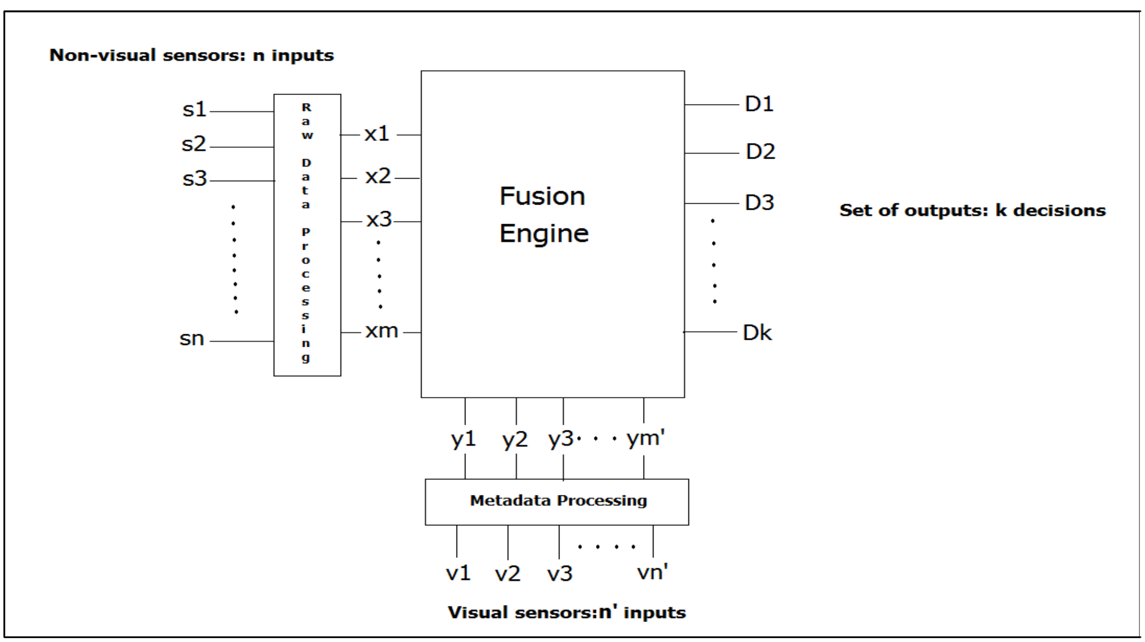


Figure 41: Fusion Engine Design

In the fusion engine design shown in Figure 41, v_1, v_2, v_3…, $v_{n'}$,represent the information collected (metadata) from the every visual sensor (n' visual sensors). These inputs are processed (metadata processing in Figure 41) to create appropriate input format. The resulting outputs of the metadata processing are also in the form of correlated information. In other words, some visual sensors can be correlated in the sense that only one output can be retrieved from them. This correlation of visual sensors results in independent inputs labeled as y_1, y_2, y_3…, $y_{m'}$ in Figure 41.

After tracking moving objects on a given video, more work is done on detecting the different features of these moving objects. Features such as the number of moving objects, the nature of the moving objects (human, animal…), and the nature of movements (fast, slow…) the objects perform are examples of information we want to feed to the fusion engine. After extracting such important information (metadata), we perform another processing on the metadata to come up with an input format compatible with the data fusion model we are using (Bayesian model).

In data fusion context, the outputs of such a model are in the form of decisions that should be performed to better serve the environment where the different types of sensors are used. As Figure 41 shows, the set of decisions D1, D2 , Dk are the independent fusion engine outputs (or decisions). These decisions can help in saving energy, restricting security, launching rescue operations and many more. Depending on what type of sensors we use, a set of relevant and efficient decisions can be formed.

4.1.2. Dempster-Schafer-Based Data Fusion in Smart Environments

Uncertainty management stems from classical methods that represent uncertainty in measurements using the Bayesian probability model to express the degree of belief in each hypothesis as a probability. The hypothesis must be mutually exclusive and this requires that all hypotheses must form a complete set of possibilities and the probabilities must sum to one. Because the Bayesian model cannot represent uncertainty along with the fact that probabilities must be assigned to each hypothesis, Dempster-Shafer introduced the concept of probability intervals to provide means to express uncertainty and that is another reason why this model is preferred in situations where probabilities cannot be assigned accurately. Other heuristic models and fuzzy calculus have also been applied to uncertainty representation for fusion applications [2].

Dempster-Shafer Theory: Dempster-Shafer theory is considered to be a generalization of the Bayesian theory of subjective probability. Dempster-Shafer allows us to "base degrees of belief for one question on probabilities for a related question" [6]. In fact, the Dempster-Shafer theory is based on two ideas: the degrees of belief for a question are obtained from subjective probabilities associated with a related question and the degrees of belief are combined using Dempster's rule "when they are based on independent items of evidence." One of the most important advantages of the Dempster-Shafer theory is that it does not associate probabilities to questions of interest as Bayesian methods do. Instead, the belief for one question is based on probabilities for a related question; therefore, the Dempster-Shafer theory can effectively model uncertainty. Additionally, the Dempster-Shafer theory doesn't require or demand the use of probabilities whenever possible [1]. Furthermore, Dempster-Shafer allows the computation of additional support and plausibility, as opposed to the Bayesian theory [2].

Dempster-Shafer model of combination evidence integrates data, independently, from r sensors' inputs in the following pattern:

$$m^{1,2,\dots,r}(C) = \frac{\sum_{c1 \cap c2 \cap \dots \cap cr \neq \emptyset} m_1(c_1), m_2(c_2), \dots, m_r(c_r)}{1 - \sum_{c1 \cap c2 \cap \dots \cap cr = \emptyset} m_1(c_1), m_2(c_2), \dots, m_r(c_r)}$$

(1)

Where

C is the proposition, so we have in this case C = $\{c_1, c_2, \dots, c_r\}$ (universal set → C),
$c_1, c_2, \dots, c_r$ are independent decisions by sensors regarding proposition C, and
$m_1, m_2, \dots, m_r$ are independent belief or mass functions.

Also, the term $\sum_{c1 \cap c2 \cap \dots \cap cr = \emptyset} m_1(c_1), m_2(c_2), \dots, m_r(c_r)$ in Equation (1) accounts for conflicts in the belief distributions from the sensors and assures that the combined belief is normalized to the unit interval.

However, Dempster-Shafer theory does not consider the quality of data that has been fused. As it is shown from equation one, only mass functions of each sensor that are fused without taking into consideration to what degree we trust each sensor in our network. Therefore, we noticed the need to a better version of Dempster-Shafer model that will take into account the weight or the confidence value of each sensor in the network (how much we trust the sensor's input). In the next section, we discuss our weighted Dempster-Shafer algorithm and present an experiment where we implemented our algorithm and compare the results to the Dempster-Shafer model.

Dynamic Weighted Dempster-Shafer**:** In this section, we suggest a method where we assign weight or confidence to every sensor involved in a given wireless sensor network. The algorithm suggests a dynamic assignment of weights to the sensors rather than a static one. The initial step, however, is a learning step where convenient weights are assigned to the sensors before they can be used in the data fusion model, which is the Dempster-Shafer in this case. The learning step is critical because we don't want the sensors to be equally trusted, but rather, we want to build a confidence of a sensor based on how well it does in terms of detections and measurements of changes in a given environment. The most important part about this algorithm is that the weights of sensors will keep updating even after the learning has been done. A description of our method is as follows.

At the very beginning, all the sensors are trusted and assigned a similar weight. A weighted Dempster-Shafer fusion of information follows where weighted mass functions that correspond to every sensor are fused. The fused value is then compared to every sensor's mass function and a new weight is computed, based on the difference between a sensor's mass and the fused mass, and assigned to each sensor. A weighted Demspter-Shafer fusion model is a model that simply multiplies every sensor's mass function by its equivalent confidence or weight before the mass can be used in the fusion process as follows:

$$m^{1,2,\dots r}(C) = \frac{\sum_{c1 \cap c2 \dots \cap cr \neq \emptyset} w_1 \cdot m_1(c_1), w_2 \cdot m_2(c_2), \dots w_r \cdot m_r(c_r)}{1 - \sum_{c1 \cap c2 \dots \cap cr \neq \emptyset} w_1 \cdot m_1(c_1), w_2 \cdot m_2(c_2), \dots w_r \cdot m_r(c_r)} \quad (2)$$

The weighted Dempster-Shafer method results in more accurate information about the likelihood of a given event since every sensor has its own weight and would contribute in the fusion based on that particular weight.

To update the weight, we multiply the initial weight of the sensor by K as follows:

$$\mathrm{w}' = \mathrm{K} \cdot \mathrm{w} \quad (3)$$

Where

w′: Sensor's new computed weight
w: Sensor's old weight
K: 1-C/2 where C is the difference between the sensor and the fused masses

The learning algorithm steps for r sensors are as follows

- $w_i = 1$ for each i=1, 2, …, r
- Apply Equation (2)
- $\Delta m_i = m^{1,2,\ldots,r}(C) - m_i(c_i)$ for each i=1, 2, …, r
- Apply Equation (3): $w_i' = \left(1 - \frac{\Delta m_i}{2}\right) \cdot w_i$ for each i=1, 2, …, r

The number of times this learning algorithm should be run depends on the application in question and can experimentally be suggested. In our analysis, we chose to run the algorithm five times before the sensors can be used in a real time environment. Table 7 below suggests mass functions from two sensors that capture changes in temperature.

Table 7: Sensors' Mass Functions

Events	Sensor 1 (s1)	Sensor 2 (s2)
Hot	30%	20%
Cold	20%	50%
Unknown	100%	100%

Using the information from Table 7, we run the algorithm five times in order to assign accurate weights to the two sensors, and the final weights are w_{s1} = 0.76 and w_{s2}= 0.60, respectively.

4.1.3. Design Optimization in Smart Environments

In this section, we discuss a key point that should be taken into consideration when building a smart environment: optimized design (placement) of sensor and networked nodes. Sensor placement is very crucial because it influences the resource management and the type of back- end processing and exploitation that must be carried out with sensed data in distributed sensor networks [9]. The main issue here is to know where exactly these sensors need to be placed and how many sensors are needed for the optimum network performance and the cost of the system.

In outdoor applications such as agricultural irrigation, sensor placement needs to be done carefully in order to optimize the sensor resources and costs. In indoor applications as well, intelligent sensor placement facilitates the unified design and operation of sensor/exploitation systems and decreases the need for excessive network communication for surveillance, target location and tracking. In fact, the use of sensors should take into consideration any obstacles that might interfere with the line of vision for IR sensors. These obstacles range from buildings to trees to uneven surfaces [9].

Any approach for such an optimization should minimize the number of sensors used in the distributed network as well as decreasing the costs and optimize the amount of data that is transferred in the network. Optimized sensor placement ensures that the resulting data contains sufficient information for the data processing center to make the decisions with sufficient data. It is discussed in [10] that there exists a close resemblance between the sensor placement problem and the placement of guards in the well-studied Art Gallery Problem (AGP) addressed by the art gallery theorem. Basically, the AGP problem deals with determination of the minimum number of guards required to cover the interior of an art gallery where the interior of the art gallery is presented by a polygon. Additionally, the sensor placement problem for target location is also closely related to the alarm placement problem. This problem deals with the placement of alarms on the nodes of a specific graph such that a single fault in the system (corresponding to a single faulty node in the graph) can be diagnosed. Furthermore, integer linear programming approach was also used to solve the problem of sensor placement on two and three dimensional grids. However, this approach has two main drawbacks: the complexity of computations makes it less appropriate for large problems and where the sensors are expected to be perfect and need to yield a binary yes/no detection in each case [9].

In our optimization approach, we implemented a design optimization tool as shown in Figure 42. This tool visualizes the outcome of the optimization solution and allows the user to modify the layout manually. The tool also helps the user optimize the placement of four types of sensors commonly used in the irrigation application, chosen here as an example. These sensor types are: soil moisture sensor, temperature sensor, wind sensor, and the carbon dioxide detector. The network also includes several actuators such as valves or gates that can impact the sensors' readings and with their decisions. The design tool facilitates the building of the best possible virtual radio frequency (RF) mesh network that will help optimize the number and location of the sensors needed and the cost of the actual implementation of such network. There are many factors that should be taken into consideration when placing sensors in an agricultural field: distance as well as obstacles between sensors or nodes in our network and the type of sensors that will be used in each case scenario. Since in our study we are considering RF network, the distance between each node in our network is very critical. As Figure 42 demonstrates, even if our network is a mesh network, we notice that some sensors are not connected either because are placed far away from one another or there are obstacles that block the communication gate. The actuators are the control units of the network that facilitate the communication between the sensors in the network as well as make the decisions needed when action is needed. For example, in case a soil moisture sensor indicates that a specific zone is dry; the actuator actually controls the valve that allows water flow to that zone.

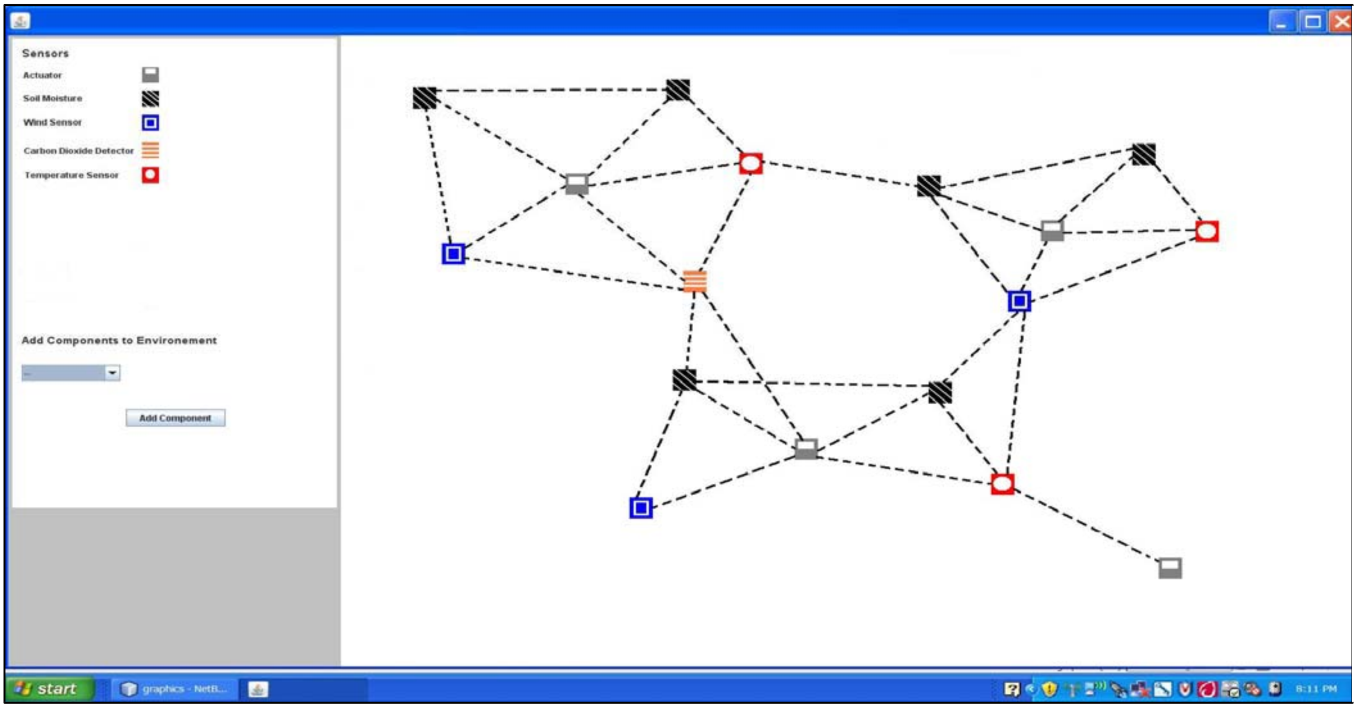


Figure 42: Smart Environment Design Tool with a Wireless Sensor Network for Irrigation

4.2 Results and Discussion

In our study of multi-sensor data fusion, we implement a simulation tool that helps us construct a virtual smart environment. The smart environment has basically different types of sensors such as: motion detector, smoke detector, daylight sensor, and other types of sensors. In addition to sensors, there are objects that can be moving around to generate case scenarios where motion is a factor to be considered. Emergency cases such as fire or flood can be studied using the implemented simulation tool. This tool is implemented using Java and it facilitates the study of multiple scenarios because the user can chose any type of sensors implemented in the tool as well as manage the environment's state such as increasing the temperature (fire case) or adding moving objects or water (flood scenario). Visual sensors are placed on the simulation grid at specific grid locations. A specific set of attributes must be defined for each sensor. These may include range, angle, sensitivity, and direction. Every sensor has a detection area and detection occurs when the coverage area and attributes of a given object overlap with the detection range and sensitivities of a given sensor. The simulation tool is our main data generator where sensors' flags and data are fed to the fusion engine where decision making process takes place.

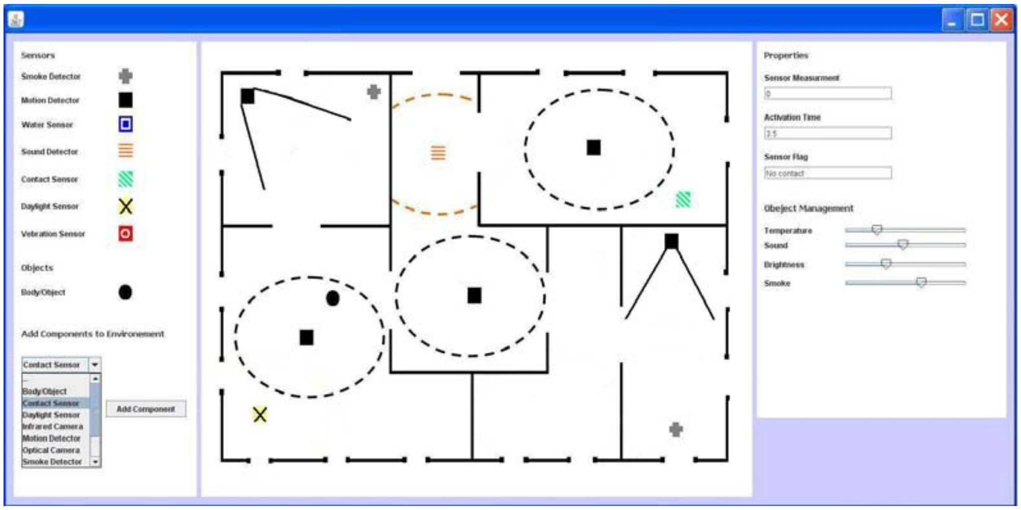


Figure 43: Simulation Tool Interface

In order to develop a reasonable method for finding a likelihood function at a given moment, we have carefully studied the behavior of the moving objects. We have conducted ten experiments where we tracked one object in every video and recorded the corresponding data. Because of space limitations, we present only the conclusions we have derived from the analysis of data. Through analyzing the graphs from the experiments, we take into consideration the factor of persistence, which merely means for how long the object(s) is moving. In order to do this, we choose a time instance from the plot and study the behavior of the moving object in previous time instants.

The probability value that will be used by the fusion engine at time t_i is computed as follows:

$y_i = \sum_{i=0}^{m-1} X(t_i) \cdot \left[\frac{1-i}{10}\right]$ where $\mathrm{X(t_i)}$ is $\mathrm{t_i}$'s equivalent area percentage

In order to find the reasonable number of previous time instants that should be included in the computation of a given likelihood function at a given instant, we further analyze the data collected from the ten experiments. We apply the formula above at t_9 for every experiment and find the equivalent likelihood function (y_9) taking into consideration m=10, 8, 6, and 4 previous time instants. Table 8 summarizes the analysis:

Table 8: Computed Likelihood Function at t_9 Using the Weighted Method

	Exp1	Exp2	Exp3	Exp4	Exp5	Exp6	Exp7	Exp8	Exp9	Exp10
m=10	117.70%	108.40%	75.83%	210.48%	62.10%	27.29%	84.31%	195.53%	72.45%	80.17%
m=8	77.50%	77.84%	55.21%	151.95%	40.81%	20.20%	58.00%	127.93%	48.65%	51.80%
m=6	56.93%	48.38%	35.18%	95.94%	22.13%	13.10%	34.40%	76.18%	28.03%	28.60%
m=4	28.60%	24.66%	18.76%	52.01%	9.48%	6.52%	16.17%	36.50%	13.81%	12.82%

From the table above, we conclude that looking back at eight or six time instants usually result in a reasonable value that gives us an idea about how intense the motion is in a given room and can safely be fed to the fusion engine. Also, the computation of a likelihood function for m=8 or 6 is easy and quicker than m=10 or more; it also doesn't take into consideration the percentage value at t_0 where usually no motion is recorded.

4.2.1. Results for Dempster-Schafer Algorithm

As mentioned above, the updating of the sensors' weight does not stop at the learning stage, but rather, weights are updated in practice as well to compensate for any sensor's failure or misreading. The table below illustrates a comparison between the quality of fused information from a regular Dempster-Shafer fusion algorithm (no weights taken into consideration) and a weighted Dempster-Shafer algorithm where the weights are updated using our suggested algorithm.

The tables below show the fusions of mass functions using both regular and weighted Dempster-Shafer methods. The dynamic weight assignment is performed throughout the process to improve the quality of the fusion. Both fusions are performed on experimental sensors' masses (four simulated data sets) that concern the frame of *discernment* (hot, cold, and unknown). In fact, we prepare the environment where the sensors' observations take place. In other words, we control and prepare the temperature parameters of the environment and then use the sensors to capture their individual reports. In this sense, we would already have an idea about how cold or hot the environment is before we rely on the sensors' observations. We then use the sensors' masses or beliefs in both fusion types (regular and weighted Dempster-Shafer).

Table 9: Fusion of Sensors' Masses for "Hot" Event

Data Sets	Dempster-Shafer Fusion	Weighted Dempster-Shafer Fusion
#1	53.55%	23.82%
#2	81.98%	35.44%
#3	64.27%	27.55%
#4	93.22%	38.53%

Table 10: Fusion of Sensors' Masses for "Cold" Event

Data Sets	Dempster-Shafer Fusion	Weighted Dempster-Shafer Fusion
#1	29.70%	13.21%
#2	26.21%	11.36%
#3	20.96%	9.12%
#4	20.14%	8.44%

As Table 10 shows, we can get better or more accurate information about an event using the weighted Dempster-Shafer method based on the pre-known information we have from the controlled parameters. For example, the first data set resulted in 53.55% "Hot" environment from the regular Dempster-Shafer versus only 23.82% from the weighted Dempster-Shafer. In fact, 53.55% was a much higher percentage to be assigned to the environmental temperature we prepared, which was not hot. The same analysis is true regarding the remaining data sets on both tables where we kept changing the parameters and capturing the differences.

Indeed, a regular Dempster-Shafer fusion cannot be trusted completely since both sensors are trusted, which means that both sensors contribute equally in reporting a change in the environment (high temperature in this example). The weighted Dempster-Shafer fusion, however, use the most updated weights in order to report the information about an event. In other terms, the confidence of every sensor is updated depending on the accuracy of the individual information it reports.

4.2.3. Smart Environment Applications

There are many applications where an optimized smart environment can improve the effectiveness and efficiency in a particular context. In this section, we identify several such applications and discuss the benefits a smart environment could provide with the tools we have covered in the previous section.

Energy Efficiency: Energy efficiency has become a real concern in this millennium. With the high technology and the new inventions, the need of energy augments to a stage where it is a necessity to manage the energy usage in order to prevent possible losses and costs. In buildings equipped with smart features, energy efficiency has been a significant benefit. In order to assist the energy saving process, most of smart buildings are equipped with day light sensors that are an innovative energy saving device. It detects an influx of daylight, and in turn automatically dims a fluorescent luminary, or series of luminaries. Daylight sensor detects any kind of light and can be used to adjust the lighting in the room to meet the needs of the room occupants.

In addition, occupancy sensors can be used in smart homes/buildings to control the usage of energy for both lighting and heating/cooling areas. For example, the rooms are illuminated or heated/cooled unless they are occupied; the occupancy sensor detects the human presence and automatically turns on the light or heating in a room. For such a system to work efficiently a multitude of sensors need to be placed in various locations in a building and the inputs from multiple sensors are used make decisions at various actuator points. The interactions and relationship between these inputs and decisions are modeled using the data fusion model in Figure 41 and the network is optimized using the tools in previous section.

Surveillance: Since smart environments are always equipped with multiple sensors and processing mechanisms, they benefit significantly most of the surveillance applications. It is obvious that when fusing data from different sources, one gets a better idea about the environment and that will enhance the decision making in case of events of interest such as a robbery or fire. In our previous work on data fusion in smart environments, we presented methods for integrating surveillance camera data with data from different sensors types' in order to detect the occupancy (as well as the number of people, suspicious presence etc.) of rooms or buildings. The control of the environment, that smart environments provide, can also be very beneficial in critical situations such as fire. In such cases, the information gathered from sensors provides accurate data about where the fire spots are, which can facilitate the evacuation operation.

Irrigation: Irrigation is another important application area where smart systems have improved water and money savings. Dukes explains that irrigation controllers that have been in use since the early 2000's are smart controllers that effectively reduce outdoor water use through monitoring site conditions such as soil moisture, plant type, or wind, and irrigating based on those parameters. In addition, these smart irrigation controllers receive feedback from the irrigated system and schedule the irrigation duration or frequency accordingly. An example that explains how water and money can be saved would be increasing watering the soil in hot or dry seasons and reducing it during cooler seasons. Generally, there are two types of smart controllers: climatologically-based controllers and soil moisture-based controllers [7].

Climatologically-based controllers are also known as Evapotranspiration (ET) controllers. In fact, ET is the process of transpiration by plants combined with evaporation that occurs from plant and soil surfaces. In general, three types of ET controllers are distinguished: signal based, historical ET, and on-site weather measurement. Signal-based ET controllers receive meteorological data from public sources or weather stations. An ET value is then calculated for a hypothetical grass surface for that site and sent to the surrounding controllers via wireless communication. The ET controller adjusts the irrigation times or days according to the climate throughout the year. The on-site weather measurement approach, on the other hand, makes use of measured weather data at the controller to calculate ET in a continuous manner and adjust the irrigation times according to the weather conditions [3].

Alternatively, soil moisture sensor controllers make use of two control strategies: "bypass" and "on-demand." The "bypass" strategy is widely used in small sites especially residential sites. In fact, the bypass soil moisture sensor controller includes a soil moisture threshold adjustment (dry to wet) that can be used to increase or decrease the sensitivity or the point at which irrigation is needed. If the current soil moisture content exceeds the threshold, this controller delays the timed irrigation. Usually, only one soil moisture sensor is used, which requires the sensor to be placed in the driest area and adjust the run times for other areas to avoid over-watering. The on-demand soil moisture sensor controller, however, starts the irrigation at a pre-programmed low soil moisture threshold and terminates irrigation at a high threshold. This type of controllers is often used in sites that involve many irrigation zones; therefore, it initiates and terminates irrigation run times in contrast to the bypass configuration that only allows irrigation events [7].

The smart environment design and simulation tools and the theoretical models we have discussed in this paper can greatly benefit irrigation applications. In most situations, at design time, the placement of the zones and the sensors within zones is an open question. Optimized sensor network design ensures the proper network operation and that the goals in the application, such as maintaining the soil moisture level, is achieved. Data fusion methods help utilize integrated and efficient processing of sensor information and better decisions to be made as a result.

4.3 Conclusions

We have demonstrated ways to use Bayesian data fusion technique in a smart environment with a heterogeneous, inter-dependent set of sensors. This was done by generating statistically independent inputs for the Bayesian fusion model and demonstrating the effect through a simulation tool. The *Dempster-Shafer Theory* is considered to be a generalization of the Bayesian theory of subjective probability. Dempster-Shafer allows us to "base degrees of belief for one question on probabilities for a related question" [6]. One of the most important advantages of the Dempster-Shafer theory is that it does not associate probabilities to questions of interest as Bayesian methods do. Instead, the belief for one question is based on probabilities for a related question; therefore, the Dempster-Shafer theory can effectively model uncertainty.

As a next step, we plan to build a Dempster-Shafer model and draw comparisons with the Bayesian model. Additionally, further experimentation is underway using a test bed created by ZigBee-based sensors that implement the smart environment and by optical and infrared cameras.

4.4 Introduction Year Two Work

The second year of work on IQ Tools for Year Two persistent Surveillance Datasets extends research done during the first year. Track Three Research was extended to set up smart building components in the field and conduct and analyze field experiments.

Smart environment is a term that refers to "a physical world that is richly and invisibly interwoven with sensors, actuators, displays, and computational elements, embedded seamlessly in the everyday objects of our lives, and connected through a continuous network." The main goal in a smart environment is to enhance the experience of individuals by replacing physical labor or repetitive tasks with automated agents.

Smart environments have many features such as remote control of devices, device communication, information acquisition and dissemination from sensor networks, enhanced services by intelligent devices, and predictive and decision-making capabilities. Technologies used in smart environments involve wireless communication, adaptive control, parallel processing, image processing, image recognition, signal prediction and classification, sensor design, motion detection, and many others.

However, the components of a smart environment do not always disseminate correct information. In other words, they do not always capture and report the real characteristics of the environment where they are placed. Problems such as missing data, wrong measurements, inconsistent readings, and incomplete data are all examples of scenarios that are likely to happen in a smart environment. In this sense, it may not be possible to rely on this kind of data when making decisions about a certain event. Therefore, it is not advisable to use data straight from the sensors without preprocessing and analysis stages. This important point has triggered the attention to investigate the preprocessing and analysis of sensory data before employing it in further decision making processes. In fact, there are many methods to detect errors in data reported by sensors in addition to those that would also correct the errors and predict the right measurement for a given event.

One of the most commonly used methods to address sensing errors is the outlier detection where readings are compared and analyzed to identify those that are distant from the rest. The after-deployment calibration is another way to deal with this issue such as the development of a mapping function that maps erroneous readings to correct ones. The parameters of the function can be obtained in many ways, but assumptions about the sensing model, dense deployment, similarity of readings among neighbors, and availability of ground truth result are usually required. The authors in [1]. suggest a new method to target data faults called FIND. This method is a sequence-based detection approach that assumes no distribution of readings. Since no distribution of readings is assumed, FIND accomplishes the detection through identifying ranking violations in node sequences, where a sequence is obtained through ordering the identifiers (IDs) of the nodes according to their readings for a given event.

In [2], the authors suggest an algorithm that uses data predictions to filter out errors caused by soft failures (failures caused by a deviation from the normal behavior). The suggested algorithm allows for a delayed reporting of data, which helps them use the observed values in the next samples and find the correct choice of value that lies between the predicted and the observed value. All error corrections are carried out at the receiver that is less resource-constrained than the sensor nodes. The framework of this work is composed of three main processes. The first process is a model of data generation that is constructed by identifying the correlations observed in sample of sensory data. The second process is another model that is used for online prediction of data and the third process is a correction block that uses the prediction history in order to correct errors detected in data.

in this project, detecting sensor data errors is accomplished through clustering the data collected from the sensors (outlier detection) using the k-medoids algorithm that exists on Wikato Environment for Knowledge Analysis (WEKA) libraries. After the detection phase, the correction phase takes place using the RBF Network algorithm that is also implemented on WEKA libraries. Both algorithms are further explained under the “suggested approach” sub-section.

4.5 Background

4.5.1. Wireless Sensor Networks

Wireless sensor networks are a type of networked systems that are characterized by "severely constrained computational and energy resources and an ad hoc operational environment" [4]. The word "wireless" means that all networking transactions and movements are accomplished without the use of wires connecting the different components of the network. The sensor networks, on the other hand, refer to a heterogeneous system that combines small sensors and actuators with general-purpose computing elements [4]. Wireless sensor networks usually can include up to thousands of "self-organizing, low power, low cost" nodes used all together to monitor a given environment where they operate. There are a lot of applications where wireless sensor networks can be very beneficial and they include: inventory control, burglar alarms, emergency response, agricultural irrigation, and military tracking. There are a lot of manufacturers of sensors and a lot of industry standards such as ZigBee that make it possible to collect and aggregate data from multiple heterogeneous sensors efficiently (mesh network topology). In this project, the Cricket system was used to establish a wireless sensor network where the nodes communication follows a star topology rather.

4.5.2. Cricket System

In order to test and improve methods that address the accuracy of information in a wireless sensor network, we have used a Cricket location system mounted at Tec^Edge in Dayton, OH. Tec^Edge Innovation and Collaboration Center is a facility owned by the Wright Brothers Institute in Dayton, OH. It is a research environment where students from different universities and high schools get together to work on challenging projects and do research in different fields. We had the opportunity to work at Tec^Edge and established the Cricket system over there for research purposes.

The Cricket system is a system that consists of a number of beacons and a listener attached on a host device. Both beacons and listeners are similar motes, and they only need to be configured either way. The way the Cricket works is simple: the beacons periodically broadcasts their space identifiers and position coordinates on a radio frequency channel that can be received by the listener. For more information about the Cricket system, users/readers can access a detailed Cricket manual on their website that is hosted by the Massachusetts Institute of Technology (MIT) [3].

The analysis of data received by the listener is the focus of this work in order to detect/correct errors existing in the collected data. Data obtained from the listener can be processed on Linux computers through the use of "cricketed," which is a daemon used to access the command interface over the network and allows for the processing of information to get the different location properties [3].

Figure 44: Snapshot of the Cricket System at Tec^Edge in Dayton, OH

The data received from the beacons is in the form of a stream where every beacon's reading can be identified through the beacon ID and the space identifier. A typical output looks like the following:

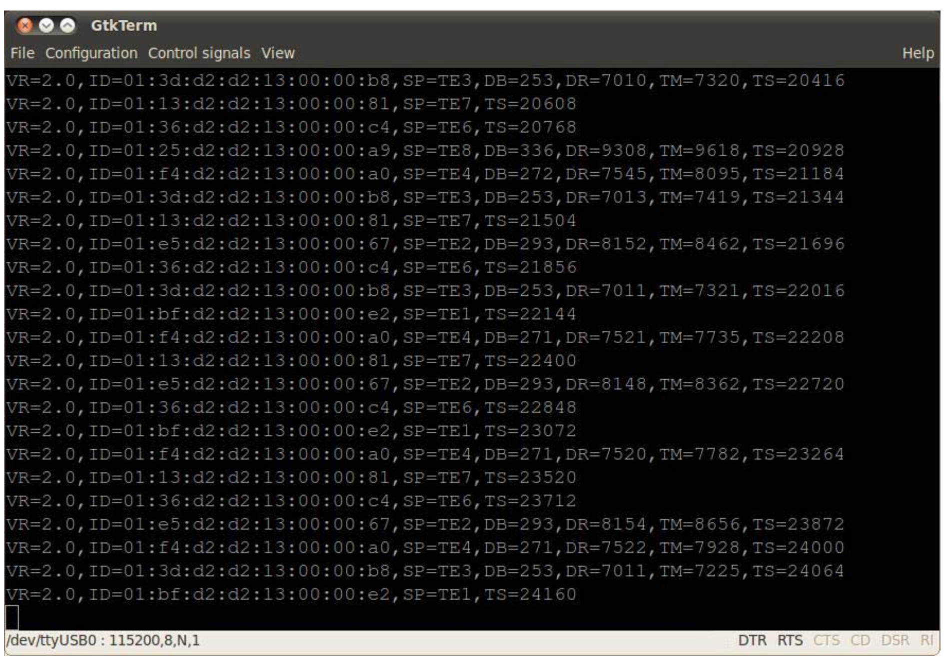


Figure 45: Snapshot of Data Stream Received by the Listener

In Figure 45, the value of the parameter ID= refers to the beacon's ID which is similar to a MAC address. This is very helpful to identify the motes that are broadcasting a message, and differentiate between the different motes. Similarly, the value of the "SP=" parameter in Figure 45 refers to the name or the space chosen for every sensor, which also helps in differentiating between information coming from different sensors. On the other hand, the value of the "DB=" parameter in Figure 45refers to the distance (information needed) between the sensor and the listener that is connected to a device capable of receiving serial data such as a computer, personal digital assistant (PDA), or a similar device. Other information can also be configured and reported by the listeners or sensors such as temperature measurements, which is the focus of this project.

4.6 Quality Dimensions

There are three important data quality dimensions that require special attention in a smart environment: *completeness*, *consistency*, and *accuracy*. Completeness refers to the degree to which a reading from a sensor is complete. In other words, the completeness dimension measures the degree to which data is missing. Consistency, on the other hand, refers to how consistent sensory data is with respect to their scheduled data transmission. Consistent sensors are sensors that report a reading every time they are used to measure the likelihood of a given event. Accuracy, which covered the most in this project, is the most interesting data quality dimension in smart environments from an information quality perspective. Accuracy implies that sensors need to report the real characteristics of the environment, which makes it a critical dimension that needs much analysis and care, especially in applications such as surveillance where it is crucial to avoid false decisions such as triggering the alarms when a false intrusion is detected.

In practice, these three dimensions are very related to each other. In other words, consistency of information is a requirement for the information to be accurate. In a similar manner, completeness of information is essential to conclude that the information is in fact accurate. In this sense, treating consistency and completeness lead to treating most of the accuracy aspects.

Tracking completeness of information requires analyzing the data stream for any missing data. In case whole information is missing, methods that use the sensor's statistical data are usually used to predict the correct value. Consistency, on the other hand, is usually caused when the sensor reports quite different readings about the same event in the same environment. This situation is usually due to sensor malfunctioning or battery and can be detected through the use of neighboring sensors' readings. Readings from neighboring sensors can effectively contribute in detecting the erroneous pattern of the sensor in question and help us build a prediction scheme to overcome the inconsistent readings. Doing this helps also the accuracy dimension where the use of neighboring sensors' readings greatly improves the accuracy and correctness of information. As a result, this part of the project is focused mainly on addressing the accuracy, which also ensures completeness and consistency of information. Ensuring completeness, consistency, and accuracy of information result in a great value added to the smart environment as a whole. *Value-added* is therefore another dimension that is enhanced in this project. The value-added dimension in the context of wireless sensor networks refers to how beneficial the sensory information is when used. In other words, "Is the data relevant in a particular environment," "Is it providing an advantage in decisions and operations?" This is in fact the definition of the value-added dimension in a smart environment.

Data quality issues are usually due to faulty nodes that manifest two types of faults: function fault and data fault. Function fault results in the crash of nodes and usually this problem is treated through using distributed approaches such as neighbor coordination or through using centralized approaches such as status updates. On the other hand, data fault implies erroneous node sensing, which leads to eventual erroneous decisions [1]. The data fault is very crucial in a lot of applications, which is the reason why this project's idea has been chosen to detect errors and correct them in a wireless sensor network.

4.7 Methods, Problem Description, and Suggested Approach

As aforementioned, the aim of this project is to improve the data quality in a smart environment through detecting errors existing in data coming from sensors and correcting those using known algorithms. The data comes from the Cricket system, which is a system usually used to locate or track a person, a robot, or simply a moving object. However, because of the lack of equipment, the Cricket system has been configured to be used as a regular temperature sensor system where every beacon is considered a temperature sensor that transmits the temperature information to the receiver. There were three experiments introduced in order to collect the data and do analysis on it. The first experiment is called the "baseline" experiment, which is an experiment where all the sensors should be operating successfully and reporting the temperature measures correctly. The second experiment introduces a gradient in temperature where a heat source blows hot air from one of the room corners. The third experiment is meant to introduce an outlier where the heat source is applied directly on one of the sensors (in this case, the targeted sensor reports an outlier measure of temperature).

The initial plan for this project was to use Kalman-Filter's two phases to detect (Phase I) errors in data and correct (Phase II) them. However, the model chosen for the data did not provide good results. Or, in other words, the results were not clear in terms of improving the current state of the data; it was not clear if the data quality had improved or worsened rather. As a result, other ways have been investigated to accomplish the two main components of the project proposal: detect errors and correct them.

4.7.1. Proposed Approach

The suggested solution for data quality improvements in a smart environment was the use of data mining algorithms in order to detect the errors in sensory data and correct them. The algorithm used to detect the errors is the k-medoids, which is similar to the k-means algorithm in concept besides the fact that k-medoids do not use the mean value of the objects in clustering. Instead, the k-medoids use the actual objects to represent the clusters using one representative object per cluster [5]. Basically, the k-medoids algorithm keeps iterating until it finally finds the most central object, called a medoid. The reason this algorithm, k-medoids, has been chosen is because the k-means algorithm is very sensitive to outliers since any object with a large value may substantially change the distribution of data.

The k-means algorithm is implemented on WEKA with choices of how the distances included in the clustering process are computed. There are two available choices: Euclidian distance and the Manhattan distance. In this project, the Manhattan distance was used in order to make the SimpleKMeans algorithm (k-means algorithm name in WEKA) compute the centroids as the component-wise median rather than mean (k-medoids approach). After clustering the temperature measures that correspond to the sensors, the sensors with measurements clustered alone under one cluster (only two clusters have been used) is investigated to verify if the sensor is reporting erroneous readings. Based on this investigation/analysis, a decision is made about the accuracy of the sensor in question.

The second part of the project, correction of the errors detected, is accomplished through the use of the radial basis function (RBF) network implemented in WEKA. The RBF network takes a nonlinear input and outputs a linear output [6]. The network is first trained using a supervised training dataset (discussed below under the application description) in order to find the RBF weights and fit the network outputs to the given inputs [6].

The RBF Network algorithm is located under the classifying algorithms in WEKA tool where the algorithm uses the k-means clustering (using the Manhattan distance in this project) to provide the basis functions in addition to learning a linear regression. The way this algorithm is used in this project is as follows: the network is first trained using a baseline dataset (data coming from the "baseline" experiment with known classes) then predict the measures for the testing data (datasets with problems or outliers). The predicted values are then used as a correction for the actual values reported by the sensor with the problem. More details are provided in the following sub-section.

4.7.2. Application Description

As it was stated in this project's proposal, the final deliverable would be an application (Java application) with a graphical user interface that allows for easy interactions with the user. The application locates and loads the data that needs analysis and outputs an evaluation of it. A correction of the sensor values or suggestions to the user is made if there is a problem with the reported data; otherwise, no corrections are suggested. The application has four buttons for different purposes, and below is screenshot of the application's graphic user interface (GUI):

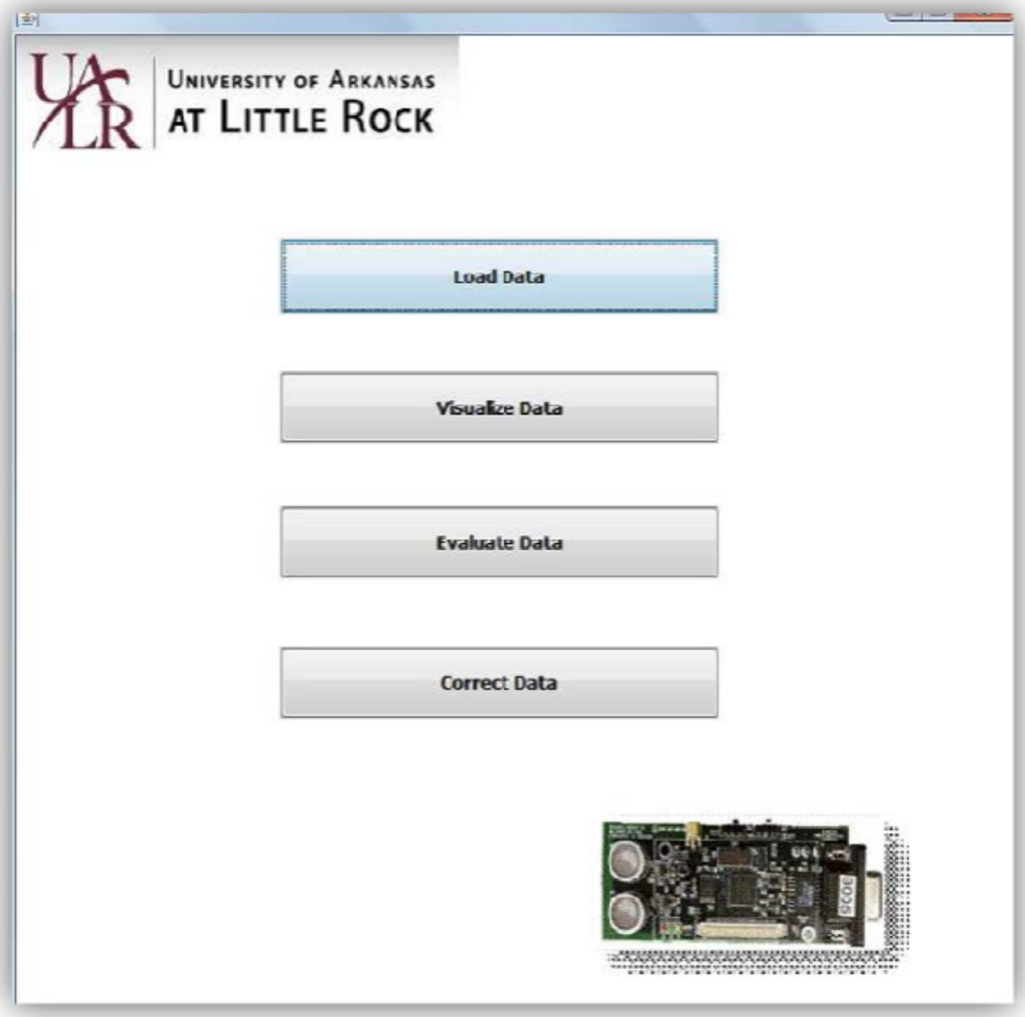


Figure 46: Graphical User Interface of the Application

Loading Data: The “Load Data” button allows the user to find the data that needs analysis. The data loaded needs to be compatible with WEKA (files have to have a “.arff” extension) so the WEKA algorithms can be applied on it. The following is a screen shot of the file finder that pops up when the button is clicked:

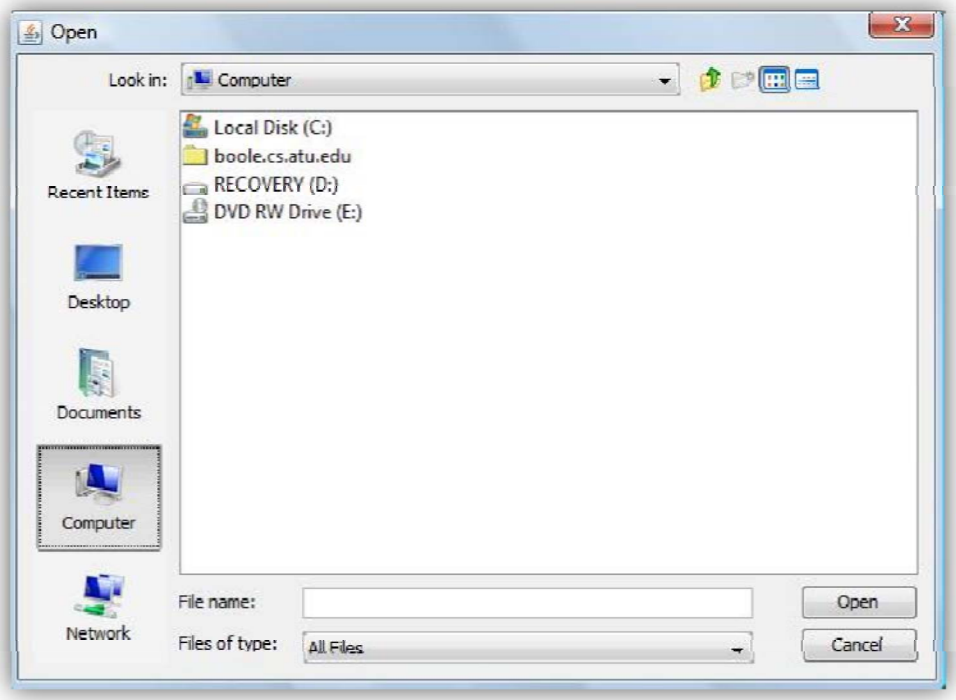


Figure 47: Locate and Load Data for Analysis

Data Visualization: The “Visualize Data” button allows for data visualization. The visualization allows the user to have an idea about the data he/she is trying to analyze. The way the data is represented is simple and user friendly. The sensor names (as they appear on the data section of the WEKA file loaded) are listed on the x-axis and the corresponding temperature measurements are represented on the y-axis. The slider labeled “jilter” allows for a better visualization of data instances since WEKA initially displays only one point that includes all readings. The following is a screen shot of how the visualization utility looks like:

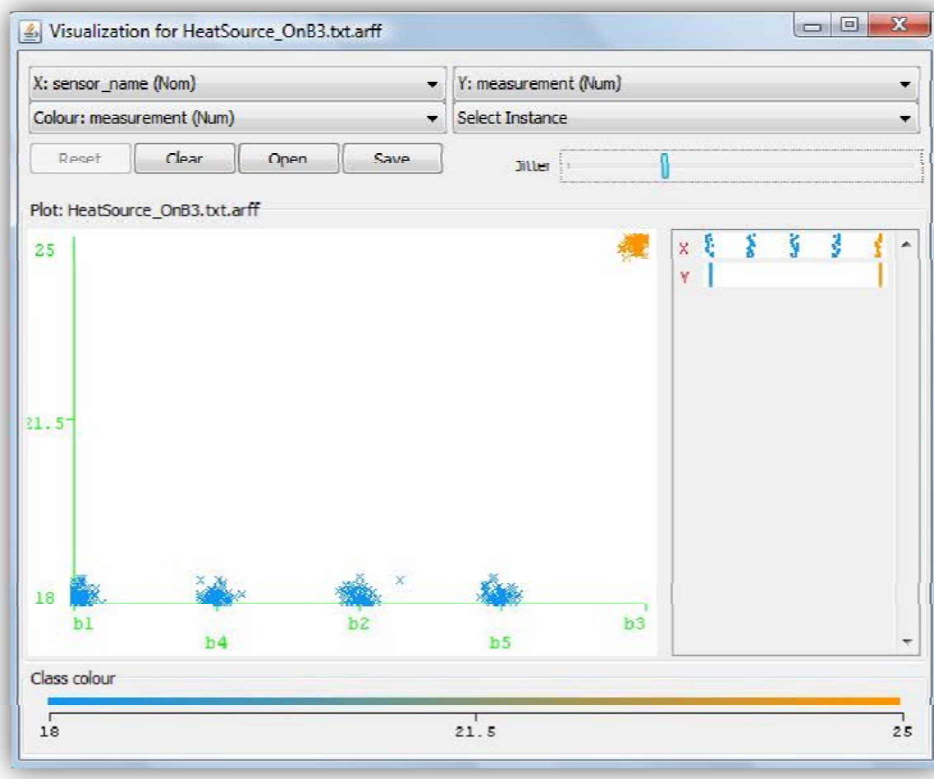


Figure 48: Data Visualization Panel

Data Evaluation: The “Evaluate Data” button performs the functionality of evaluating the loaded data and making the decision about the correctness and accuracy of sensors. This functionality is implemented using the k-medoids algorithm to cluster the data. Two clusters are used to differentiate between the sensor’s measurements that were clustered alone. In other words, the aim is to find out which sensors name its data have been clustered alone under one cluster. The result of clustering the data is a matrix that shows the number of instances that have been clustered under each cluster with their corresponding sensor names. An analysis of this matrix allows for determining if the sensor is reporting errors for real. To do this, every sensor has its location recorded. This location information represents the real physical location of every sensor at Tec^Edge in Dayton, OH. Knowing the location coordinates allow for measuring the distances between sensors which allows for neighbor identification. In this sense, if a sensor’s reported measurements are clustered alone under one single cluster, the decision that the sensor is likely to be reporting errors is not made until the neighboring sensors’ reported measurements are checked and compared to its measurements. If the sensors’ readings are way too larger or less than the neighboring sensors’ readings, the sensor is then identified as an erroneous one. An informing message is then popped up to let the user know about the sensors’ performances. Below is a screen shot of the evaluation functionality.

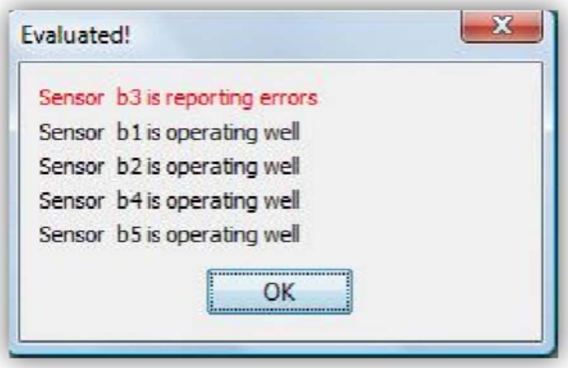


Figure 49: Evaluation Results

Data Correction: The last button labeled "Correct Data" gives the user a suggestion of the correct measurement that the sensor with problems is supposed to be reporting instead. To add this functionality, the history of every sensor has been collected and used as a training dataset that is fed to the classifier (RBF Network). The history is a dataset that has sensors' readings. This historical data has been collected in a normal environment (similar to the one where test data was collected) where the sensors are assumed to be operating correctly and accurately. The "baseline" experiment described above explains the environment where this data was collected (no heat source was introduced in the room). After training the network that is supposed to use clustering information from the first phase (evaluation) and learn a linear regression model, the testing data, which is the data that the user has loaded for evaluation, is then evaluated against the built model to suggest correct readings for the sensor in question. After the correction has been suggested, an informing message pops up for the user to read the suggested corrected value. This component, does not, however, modify the data that was loaded initially for analysis. This application is meant to evaluate the data for problems and suggest correct values rather than modify and do any changes on the data. The following is a screen shot of the message that pops up to the user when a correction is made.

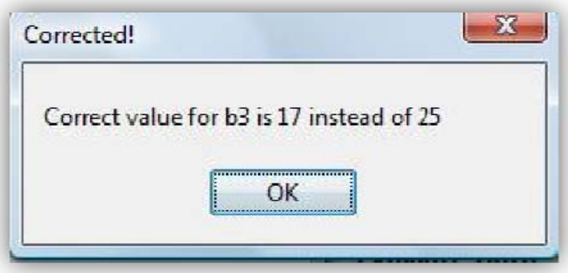


Figure 50: Correction Message

4.8 Results and Discussions

Detection and correction of sensor readings help improve accuracy and therefore completeness and consistency. Accuracy is improved in the sense that the predicted value is a better representation of the real characteristics of the environment. Detecting the erroneous reported measurement and substituting that information with a predicted value contribute in a better representation of the real characteristics of the environment where the sensors are operating. This predicted value matches and follows the same distribution as the sensors neighboring sensors' readings. In this sense, both completeness and consistency are ensured. Completeness is maintained through always ensuring a measurement value for the sensor in question (the predicted value substitutes the missing sensors' measurement). Consistency, on the other hand, is maintained through ensuring that the sensors' measurements are following the same distribution as the baseline data and that they are consistent with each other at any time a measurement is received.

The value added using this application is the ability to benefit from the wireless sensor network as a whole. In other words, if the sensors data collected from a wireless sensor network is guaranteed to be accurate, complete, and consistent, then that means the data can be used directly or indirectly in decisions without any problem. Decisions include triggering alarms (surveillance environment), irrigating the greenhouse (agricultural environment), or any other environment that needs and requires accurate sensor monitoring. The key here is when the decision is made, then it should be correct and beneficial to the application where it is used. In this project, ensuring correct and accurate temperature measurements allow for better use of this kind of information in wherever the information is needed. A good illustrating example would be monitoring and controlling a room against intrusions. If the sensors are providing good quality temperature measurements, then the change in temperature is correctly and accurately captured by the sensors and therefore, the decision to trigger burglar alarms or send the guards for support is 100% correct.

4.9 Conclusion

IQ plays a huge role in smart environments. Without sensor data pre-processing, there is always the risk of using low quality data and therefore making wrong decisions. Detecting and correcting sensory data is therefore, very important in every smart application where sensors play the main role in reporting facts about the environment. In this project, clustering temperature measurements helps identify the sensor that needs attention. Data analysis of the clustering results helps in making a decision about the accuracy or the correctness of the sensor in question and therefore, triggering the correction stage. Correcting the sensor's wrong measurement through using a predicted value ensures a better representation of the real environment and maintains complete and consistent sensors' readings. With an improved IQ in a wireless sensor environment, the sensory information is trusted and used safely in different decisions that would magnify the value added to the application where the sensors are deployed.

5.0 REFERENCES

Track One References

[1] A.Moorthy, A. Bovik, “Visual Importance Pooling for Image Quality Assessment,” *IEEE Journal of Selected Topics in Signal Processing*, **Volume 3**, No. 2, April 2009.

[2] H.R. Wu, K.R. Rao, “Digital Video Image Quality and Perceptual Coding,” *CRC Press*, 2006.

[3] Z. Wang and A. C. Bovik, "Mean Squared Error: Love It or Leave It?”, IEEE *Signal Processing Magazine,* January 2009.

[4] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity,” *IEEE Transactions on Image Processing,* **Volume 13,** No.4, pp. 600- 612, April 2004.

[5] H. R. Sheikh, M. Sabir, and A. C. Bovik, “A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms,” *IEEE Trans. Image Process.*, **Volume 15**, No. 11, pp. 3440–3451, November 2006.

[6] H.R. Sheikh.and A.C. Bovik, "Image Information and Visual Quality," *IEEE Transactions on Image Processing,* **Volume 15**, No.2, pp. 430- 444, February 2006.

[7] Q. Li, Z. Wang, “Video Quality Assessment by Incorporating a Motion Perception Model,” *ICIP’2007*, pp. 173-176.

[8] Qiang Li and Zhou Wang, “Video Quality Assessment by Incorporating a Motion Perception Model,” in Image Processing, 2007. ICIP 2007. *IEEE International Conference on 2007,* **Volume 2**, pp. 173–176.

[9] Qi Ma, Liming Zhang, “Image Quality Assessment with Visual Attention,” *ICPR 2008*, 1-4

[10] M.Milanova , S. Rubin, R. Kountchev, V. Todorov, R. Kountcheva, “Combined Visual Attention Model for Video Sequences,” *IEEE ICPR 2008*, 1-4.

[11] P. L. Callet, F. Autrusseau, ‘Subjective Quality Assessment IRCCyN/IVC Database,’ http://www.irccyn.ec-nantes.fr/ivcdb/.

[12] H.R. Sheikh, Z.Wang, L. Cormack and A.C. Bovik, "LIVE Image Quality Assessment Database Release 2,” http://live.ece.utexas.edu/research/quality.

[13] S. Winkler and P. Mohandus, The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics, IEEE Transaction on Broadcasting **Volume 54**, No. 3, pp. 600-608, 2008

[14] L. Itti C. Koch and E. Niebur, “A Model of Saliency-Based Visual Attention for Rapid Scene Analysis,” *IEEE Trans. Pattern Anal. Mach. Intel*, **Volume 20**, No.11, pp.1254-1259, November 1998.

[15] U. Rajashekar , A. Bovik and L. Cormack, Caffe, “A Gaze Attentive Fixation Finding Engine,” *IEEE Trans. Image, Processing,* **Volume 17**, No. 4, pp. 564-573, April 2008.

[16] A.Moorthy, A. Bovik, “Visual Importance Pooling for Image Quality Assessment,” *IEEE Journal of Selected Topics in Signal Processing*, **Volume 3**, No. 2, April 2009.

Track Two References

[1] Anderson, Tom H. C., Reward for Being the Top Market Researcher on Twitter. *Tom H. C. Anderson - Next Gen Market Research.* [Online] Next Gen Market Researcher, 3 February 2010 [Cited: 4 February 2010] http://www.tomhcanderson.com/2010/02/03/reward-of-being-the-top-market-researcher-on-twitter/.

[2] Morozov, Evgeny. Iran Elections: A Twitter Revolution? *The Washington Post.* [Online] June 17, 2009. [Cited: February 9, 2010.] http://www.washingtonpost.com/wp-dyn/content/discussion/2009/06/17/DI2009061702232.html.

[3] Beaumont, Claudine. Mumbai attacks: Twitter and Flickr Used to Break News. *Telegraph.co.uk.* [Online] Novomber 27, 2008. [Cited: 9 February 2010.] http://www.telegraph.co.uk/news/worldnews/asia/india/3530640/Mumbai-attacks-Twitter-and-Flickr-used-to-break-news-Bombay-India.html.

[4] (Computerworld), Heather Havenstein. LA Fire Department all 'aTwitter' over Web 2.0. *PC World.* [Online] August 3, 2007. [Cited: 9 February 2010.] http://www.pcworld.com/article/135518/la_fire_department_all_atwitter_over_web_20.html.

[5] Clint Boulton. Twitter Use Tapering Off in the U.S., but Rising Overseas - Web Services, Web 2.0 & SOA. eWeek.com. [Online] January 15, 2010. [Cited: February 28, 2010.] http://www.eweek.com/c/a/Web-Services-Web-20-and-SOA/Twitter-Use-Declining-in-the-US-But-Rising-Overseas-592323/.

[6] Tzu-Wei Hsu, et al. "MonkEllipse: Visualizing the History of Information Visualization." s.l. IEEE Symposium on Information Visualization, (r9), 2004. 10.1109/INFVIS.2004.48.

[7] Havenstein, Heather. LA Fire Department all 'aTwitter' over Web 2.0. PCWorld. [Online] August 3, 2007. [Cited: February 15, 2010.] http://www.pcworld.com/article/135518/la_fire_department_all_atwitter_over_web_20.html.

[8] Tabor, Damon. LAFD's One-Man Geek Squad Brings Web 2.0 to Firefighting. Wired Magazine. [Online] October 20, 2008. [Cited: February 15, 2010.] http://www.wired.com/entertainment/theweb/magazine/16-11/st_firefight.

[9] Data Quality in Context. Strong, Diane M., Lee, Yang W. and Wang, Richard Y. 5, s.l. : ACM Publication, May 1997, Communications of the ACM, **Volume 40**, pp. 8.

Track Three References

[1] Challa, S., Koks, D., An Introduction to Bayesian and Dempster-Shafer Data Fusion, http://robotics.caltech.edu/~jerma/research_papers/BayesChapmanKolmogorov.pdf, Accessed 28 October 2009.

[2] Waltz, Edward, James, L., *Multi-Sensor Data Fusion*. Norwood: Artech House, Inc, 1990.

[3] Fisher C., Eitel L., Richard W., and Shobha C., *Introduction to Information Quality*. Cambridge: MITIQ, 2006.

[4] H.B. Mitchell, *Multi-Sensor Data Fusion*. Berlin: Springer, 2007.

[5] Snidaro, L., Claudio, P., Christian, M., Gian, F., Visual Sensor Technology for Advanced Surveillance Systems: Historical View, Technological Aspects and Research Activities in Italy, http://www.mdpi.com/1424-8220/9/4/2252/pdf, Accessed 12 November 2009.

[6] Shafer, G., Dempster-Shafer Theory, http://www.glennshafer.com/assets/downloads/articles/article48.pdf, Accessed 28 October 2009.

[7] Dukes, Michael. "Smart Irrigation Controllers: What Makes an Irrigation Controller Smart?" University of Florida. 21 February 2010 <http://edis.ifas.ufl.edu/pdffiles/AE/AE44200.pdf>.

Appendices of Published Papers

APPENDIX 1: TRACK 1 - IDENTIFY INFORMATION REQUIREMENTS FOR THE LAYERED SENSOR DOMAIN

Appendix 1-A: "Image Quality Assessment based on Salient Region Detection."

"Image Quality Assessment based on Salient Region Detection." Engin Mendi, Mariofanna Milanova, Yinle Zhou, John Talburt—University of Arkansas at Little Rock.

Abstract *- Image quality assessment has a great importance in several image processing applications. Recently, various objective image quality metrics have been proposed in order to predict the human visual perception. In this paper, novel image quality metrics, S-SSIM (saliency-based structural similarity index) and S-VIF (saliency-based visual information fidelity), are proposed based on frequency-tuned salient region. Saliency maps are produced from the color and luminance features of the image. SSIM and VIF in pixel domain are modified by the weighting factors of the saliency maps. We validated our approach using 2 image databases as test bed: These databases contain subjective scores for each image. Our results showed that our technique is more correlated with human subjective perception.*

Image Quality Assessment Based on Salient Region Detection

Engin Mendi[1], Mariofanna Milanova[2], . Yinle Zhou[3], John . Talburt[3]
University of Arkansas at Little Rock, 2801 S. University Ave , Little Rock , Arkansas , USA
Department of Applied Science[1],
Department of Computer Science[2],
Department of Information Science[3]
esmendi@ualr.edu, mgmilanova@ualr.edu, yxzhou@ualr.edu, jrtalburt@ualr.edu

Abstract: Image quality assessment has a great importance in several image processing applications. Recently, various objective image quality metrics have been proposed in order to predict the human visual perception. In this paper, novel image quality metrics, S-SSIM (saliency-based structural similarity index) and S-VIF (saliency-based visual information fidelity), are proposed based on frequency-tuned salient region. Saliency maps are produced from the color and luminance features of the image. SSIM and VIF in pixel domain are modified by the weighting factors of the saliency maps. We validated our approach using 2 image databases as test bed: These databases contain subjective scores for each image. Our results showed that our technique is more correlated with human subjective perception.

1 Introduction

Image quality assessment has a great importance in several image and video processing applications such as filter design, image compression, restoration, denoising, reconstruction, and classification. The aim of image quality assessment is predicting image quality of display output perceived by the final user. Multimedia contents are subjected to the variety of artifacts during acquisition, processing, storage and delivering, which may lead to reductions in the quality. Our image quality assessment module dynamically monitor and adjust the image quality, so that the output quality of the image or video presented to the user can be maximized for available resources such as network conditions and bandwidth requirements.

IQA methods fall into two categories: subjective assessment by humans and objective assessment by algorithms. Subjective image quality experiments are classical statistical measurements how humans pensive the image quality. Subjective measures are determined by Mean Opinion Score (MOS) which relies on human perception.

The mathematical tools for subjective assessment of image quality are well define, but still there remain certain practical aspects how to design efficient experiment. While subjective assessment is the ultimate judge of image quality, it is time – consuming and cannot be implemented in real time quality score. This is the main reason to motivate development of algorithms which predict subjective image quality measure accurately. In [1] how “well” an algorithm performs is defined by how well it correlates with human perception of quality. Objective quality metrics are algorithms designed to characterize the quality of image and predict viewer opinion. Different types of objective metrics exist as illustrated in paper [2].

They are based on mathematical measurements which are practical to apply without need of human observers. Objective quality metrics can be classified into 3 metrics: Full Reference (FR), Reduced Reference (RR) and No Reference (NR). All

Distribution A; Approved for Public Release; Distributed Unlimited; 88ABW/PA cleared 24 September 2012 as 88ABW-2012-5092.

these metrics are based on the availability of original non-distorted reference image which will be compared with the corresponding distorted image. In FR case, reference image information is available; in RR case, partial information of reference image is known and no information about the reference image is available in the NR case.

In the image processing community more than 50 years mean squared error (MSE) are being used as quasi –standard fidelity metrics. The MSE sill continue to be widely used as a signal fidelity measure, but at the same time there are recent studies to developed more advanced signal fidelity measures, especially in applications where perceptual criteria might be relevant . It is interesting to demonstrate how the image quality is measured for different regions in an image. It is obvious that different regions in the image may not stand the same importance. Visual importance has been explored in the context of visual saliency [3], fixation calculation [1]. In [4], one experiment to record the gaze coordinates corresponding to the human eye movements and the Gaze – Attentive Fixation Finding Engine (GAFFE) was proposed. In [1] the researchers are using GAFFE to find points of potential visual importance and one algorithm for fixation-based and quality – based weighting was developed. The region-of interest based image quality assessment still remains unexplored.

In this study, we developed a novel image quality metrics, S-SSIM (saliency-based structural similarity index) and S-VIF (saliency-based visual information fidelity), based on frequency-tuned salient region detection introduced by. Saliency maps are produced from the color and luminance features of the image. Structural similarity index (SSIM) and visual information fidelity (VIF) in pixel domain are modified by the weighting factors of the saliency maps. Our results showed that our technique is more correlated with human subjective perception. The rest of this paper is organized as follows: Section 2 provides a brief overview of SSIM and VIF in pixel domain. Proposed image quality metrics based on frequency-tuned salient region are presented in Section 4. We present the results of our approach in Section 4. Finally, in Section 5 the conclusions of this paper are summarized.

2 Previous Work

2.1 SSIM

Consider two images $x = \{x_i \mid i = 1,2,...,N\}$ and $y = \{y_i \mid i = 1,2,...,N\}$ where N is the number of pixels and x_i and y_i are the i th pixels of the images of x and y , respectively. SSIM- $\mathrm{SSIM}(x,y)$ combines three comparison components, namely luminance- $l(x,y)$, contrast- $c(x,y)$ and structure- $s(x,y)$ [5]:

$$\mathrm{SSIM}(x,y) = f(l(x,y), c(x,y), s(x,y))$$

(1)

Luminance, contrast and structure comparisons are defined as follows:

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad C_1 = (K_1 L)^2$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad C_2 = (K_2 L)^2$$

(2)

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \quad C_3 = \frac{C_2}{2}$$

Distribution A; Approved for Public Release; Distributed Unlimited; 88ABW/PA cleared 24 September 2012 as 88ABW-2012-5092.

where μ_x, μ_y, σ_x, σ_y and σ_{xy} are means of x and y, variances of x and y and correlation coefficient between x and y. K_1 and K_2 are scalar constants that $K_1, K_2 << 1$ and L is the dynamic range of the pixel values. Finally, SSIM index yields to:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

(3)

2.2 VIF in Pixel Domain

VIF index relates image fidelity to the mutual information between the test and the reference images using source and distortion models and as well as human visual system model. It is given as [6]:

$$\mathrm{VIF} = \frac{\sum_{j=1}^{S}\sum_{i=1}^{M_j} I(C_{i,j}; F_{i,j})}{\sum_{j=1}^{S}\sum_{i=1}^{M_j} I(C_{i,j}; E_{i,j})}$$

(4)

$I(C_{i,j}; F_{i,j})$ and $I(C_{i,j}; E_{i,j})$ represent the information perceived by the human observer from a particular sub band in the reference and the test images respectively. C is a block vector from a given location in the reference image, E is the perception of block C by a human observer from reference image, which can be represented as $E = C + N$, where n is additive noise. F is the perception of block C by a human observer from test image, which can be represented as $E = D + N$. D is the block vector from the test image given as $D = GC + V$ where G and V are the blur and noise distortions, respectively. S denotes the number of all sub-bands and M_j is the number of blocks at j th sub-band.

3 Image Quality Assessment with Frequency-tuned Saliency Map

In recent years, it has become clear that many problems in perception organization are difficult to solve without introducing the contextual information of a visual scene. Subjects often search for the component feature of a target rather that the target itself, even if the target is a simple geometric form. Most computational models of attention ignore contextual information provided by the correlation between objects and the scene. Schyns and Oliva [7] showed that a coarse representation of the scene initiates semantic recognition before the identification of objects is processed. Many studies support the idea that scene semantics can be available early in the chain of information processing and suggest that scene recognition may not require object recognition as a first step [8].Human can recognize the scene even using low-special frequency image.

Another reason for features –driven attention is that this reflects the attempt of the eye to maximize the information it can gather at each fixation [9]. The purpose of early visual processing is to transform the highly redundant sensory input into more efficient factorial code. At the same time the human visual system has evolved multiple mechanisms for controlling gaze. Tracking can be formulated in a probabilistic framework in both the future- driven and intensity-driven settings. The principal component analysis (PCA) and the independent component analysis (ICA) are two common techniques that allow for probabilistic treatment. The PCA assumes the data distribution has a Gaussian structure and model data with an appropriate orthogonal basis functions. The ICA generalizes PCA by permitting non-Gaussian distributions and non-orthogonal bases. However, these techniques do not allow noise to be modeled separately from the signal structure, and they do not permit overcomplete codes in which there are more basis functions than input dimensions. Bell and Sejnowki [10] applied their Infomax-based ICA algorithm to image coding and reported that the independent components of the natural scenes resemble edge filters. Such Gabor-like filters are believed to be a good model of the spatiotemporal receptive fields of simples' cells in Primary visual cortex (V1). In [11] Olshausen and Field argued for maximizing the sparseness of the distribution of output activities, or "minimum entropy"coding as a good feature detector. In this study we propose to model conjunction search. Conjunction search (a search for a unique combination of two features – e.g, orientation and spatial frequency – among distractions that share only one of these features) examines how the system combines features into perceptual wholes. We propose to improve the effectiveness of the decomposition algorithm by providing the algorithm with classification awareness. Attentional guidance does not depend solely on local visual features, but must also include the effects of interactions among features. The idea is to group filters (basis components) which become responsible for extracting similar features. A certain feature will be shared by the nearest neighbors of fixations

In this study we propose visual attention model based on the extended Frequency-tuned saliency model [12] and incorporating conjunction search [9]. Saliency maps are produced from the color and luminance features of the image. Saliency map S is formulated for the image I as follows:

$$S(x,y) = \left\| I_{\mu} - I_{w_{hc}}(x,y) \right\|$$

(5)

I_{μ} is the mean image feature vector, $I_{w_{hc}}(x,y)$ is the corresponding pixel vector value of the blurred image from the original image and $\|.\|$ is the Euclidean distance. Each pixel location is the Lab color space vector, i.e. $[L,a,b]^T$.

3.1 S-SSIM and S-VIF in Pixel Domain

In human visual system, the importance of a visual event should increase with the information content, and decrease with the perceptual uncertainty [13], we incorporated saliency map as weighting function into the SSIM and VIF indexes. So saliency factors can be instated into the quality metrics. The weighting function is:

$$w(x,y) = \left\| I_{\mu} - I_{w_{hc}}(x,y) \right\| \tag{6}$$

We define saliency-based SSIM as S-SSIM and saliency-based VIF as S-VIF as follows:

$$\text{S-SSIM} = \frac{\sum_x \sum_y w(x,y)\text{SSIM}(x,y)}{\sum_x \sum_y w(x,y)}$$

$$\text{S-VIF} = \frac{\sum_x \sum_y w(C,F)\text{VIF}(C,F)}{\sum_x \sum_y w(C,F)} \tag{7}$$

SSIM and VIF in pixel domain mainly focus on local information and do not take global saliency features into consideration [14] Figure 1 shows an example case that SSIM and VIF in pixel domain fail. Figure 1(a) and Figure 1 (b) show a reference image and its frequency tuned saliency map, respectively. In Figure 1(c) and Figure 1(e), the images are distorted at both visually attended and less-attended locations by Gaussian noise and blurring effect, respectively. Less amount of same distortions are also applied to the images at only less attended locations in Figure 1(d) and Figure 1(f). It is easy to see that the quality of images in Figure 1(d) and Figure 1(f) are better than of Figure 1(c) and Figure 1(e). Even though the amounts of distortion effects are greater in Figure 1(c) and Figure 1(d), SSIM and VIF in pixel domain give incorrect results. As shown in Table 1, S-SSIM and S-VIF in pixel domain scores are more realistic.

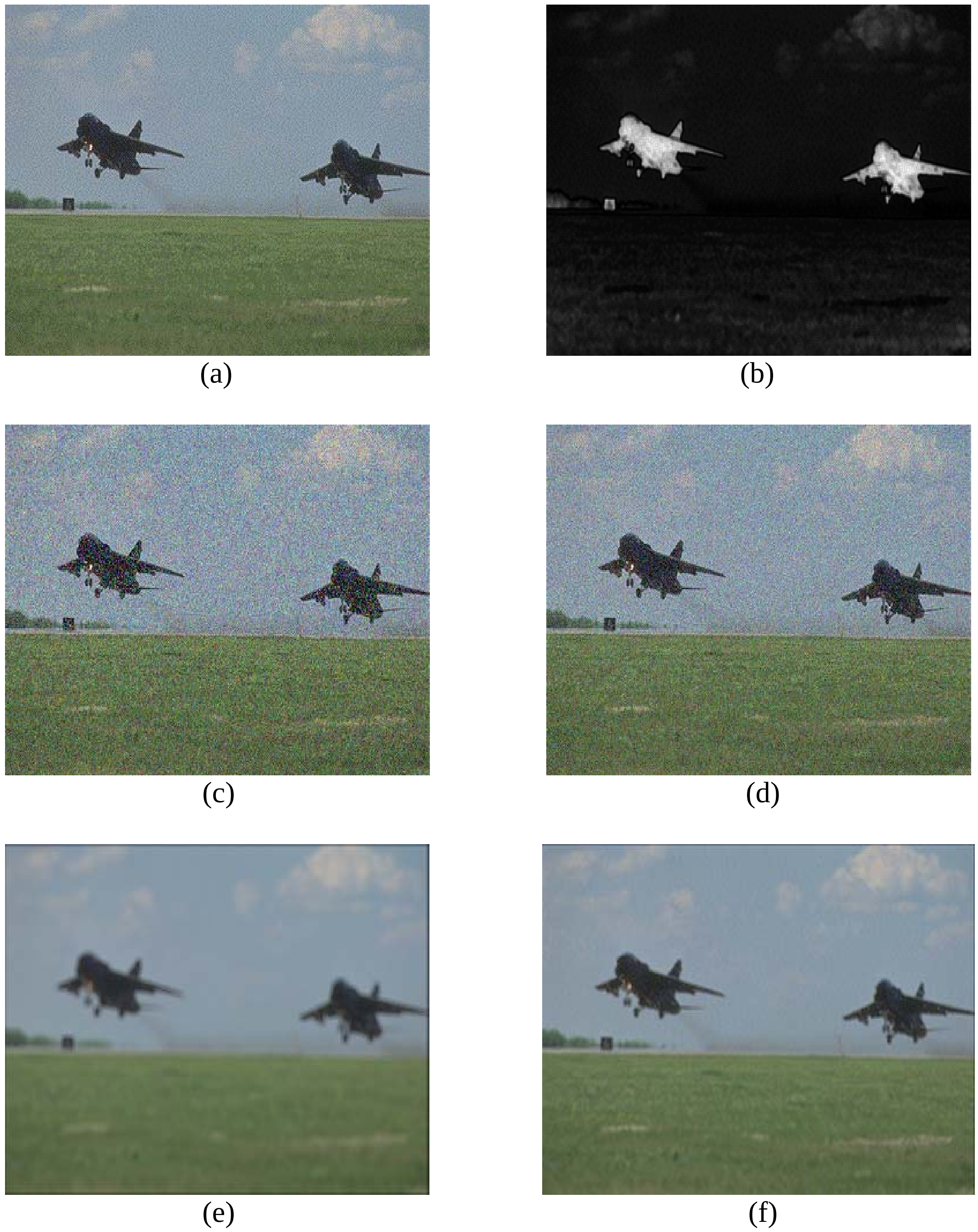


Figure 1: a) reference image, b) saliency map of the reference image), c) Distorted image with higher amount of Gaussian noise applied to attended and less-attended locations, d) Distorted image with less amount of Gaussian noise applied to only less-attended locations, c) Distorted image with higher amount of blurring effect applied to attended and less-attended locations, d) Distorted image with less amount of blurring effect applied to only less-attended locations

Table 1: Scores of SSIM, S-SSIM, VIF and S-VIF in pixel domains for images in Figure 1

	SSIM	S-SSIM	VIF in pixel	S-VIF in pixel
Fig. 2(c)	0.5976	0.8319	0.6886	0.2196
Fig. 2(d)	0.724	0.5772	0.7774	0.1449
Fig. 2(e)	0.3851	0.865	0.6751	0.3136
Fig. 2(f)	0.4463	0.6452	0.9336	0.2436

4 Experimental Results

We validated our approach using 2 image databases as test bed: These databases contain subjective scores for each image. First is the IVC Image database [15] consisting of 10 reference images with 235 distorted images (JPEG, JPEG2000, LAR coded and blurred). Second is the LIVE Image Database [16] consisting of 29 original images and 460 distorted images (227 JPEG2000 images and 233 JPEG images.) Non-linear regression analysis has been performed to fit the data. The Pearson correlation coefficient is used to measure the association between subjective and objective scores.

Figure 2 and 3 show the results for IVC and LIVE databases, respectively. Each sample point represents the subjective/objective scores of one test image. The y axis in the figures denotes the subjective scores in the databases. The x axis denotes the predicted quality of images after a nonlinear regression toward 4 objective scores, which are SSIM, S-SSIM, VIF and S-VIF in pixel domains, respectively. The Pearson validation scores between assessment metrics are depicted in Table 2.

The Pearson correlation coefficient varying from -1 to 1 is widely used to measure the association between two variables. High absolute values mean that the two variables being evaluated have high correlation. As shown in Table 2, our technique is more correlated with human subjective perception.

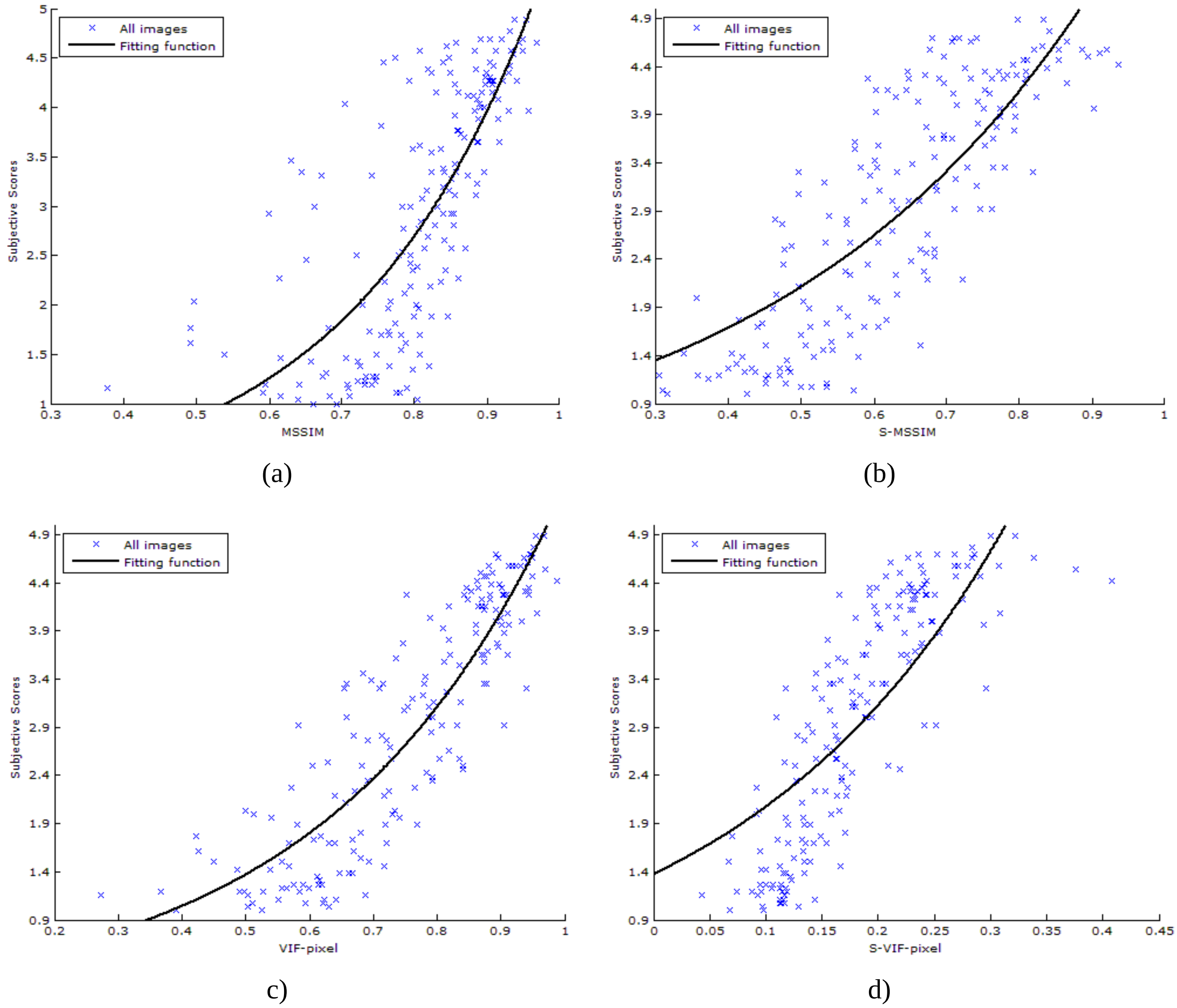


Figure 2: Scatter plots of subjective/objective scores on IVC Database. (a) SSIM; (b) S-SSIM, c) VIF in pixel domain, d) S-VIF in pixel domain

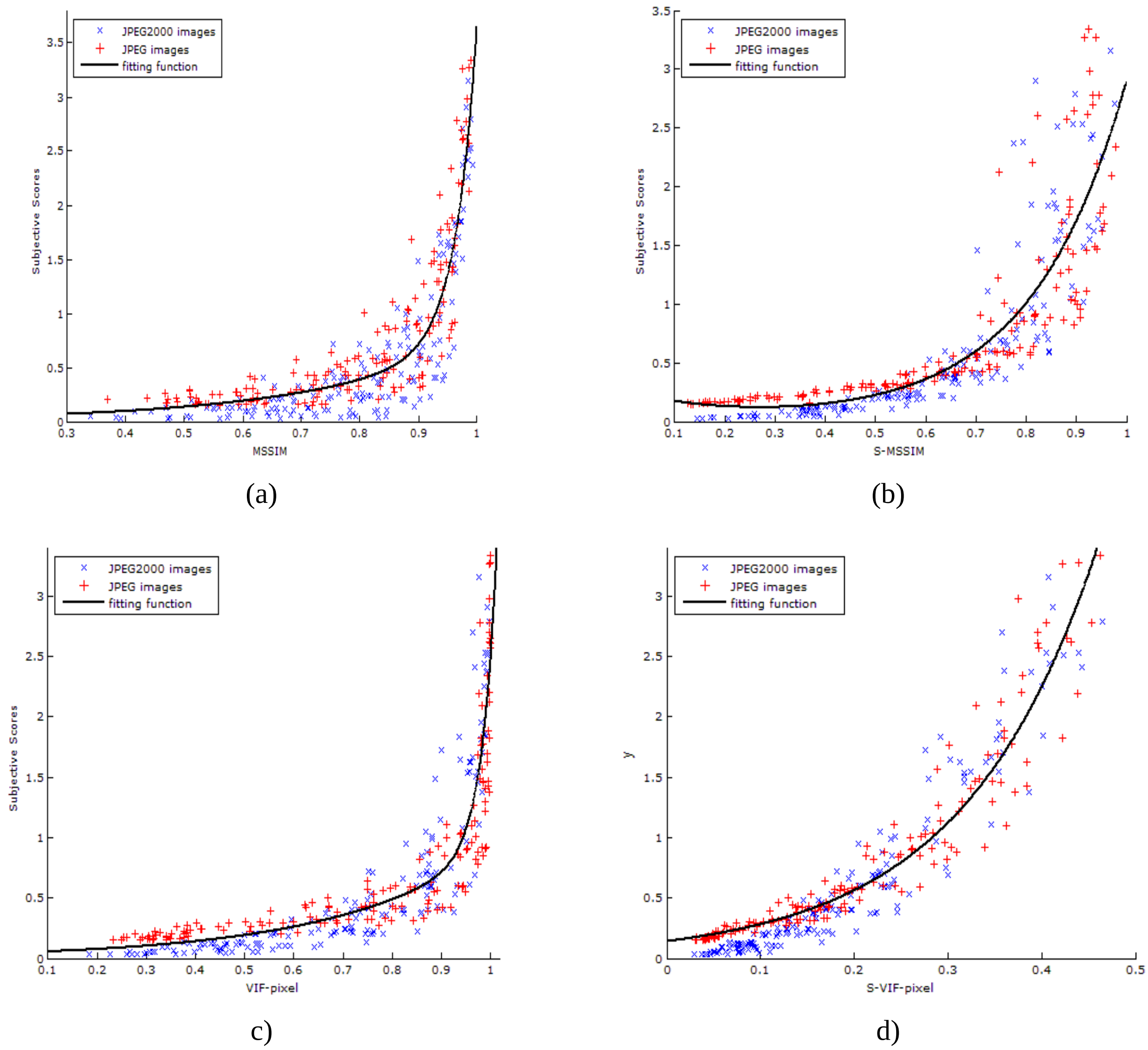


Figure 3: Scatter plots of subjective/objective scores on LIVE Database. Red points and blue points denote JPEG and JPEG2000 images, respectively, (a) SSIM; (b) S-SSIM, c) VIF in pixel domain, d) S-VIF in pixel domain

Table 2: Pearson correlation coefficients

	SSIM	S-SSIM	VIF-pixel	S-VIF-pixel
IVC - all images	0.7047	0.8261	0.8435	0.8715
LIVE - JPEG&JPEG2000 images	0.6823	0.7475	0.7126	0.9083

5 Conclusions

This paper presents two novel image quality metrics, S-SSIM and S-VIF in pixel domain. The metrics are based on frequency-tuned salient region detection and computationally inexpensive. Salient region detection captures full resolution saliency maps exploiting the color and luminance features of the images. Saliency maps are then set as weighting functions and incorporated in to SSIM and VIF in pixel domain. The approach has been validated using two image databases: 1) IVC Image database consisting of 10 reference images with 235 distorted images (JPEG, JPEG2000, LAR coded and blurred) and LIVE Image Database consisting of 29 original images and 460 distorted images (227 JPEG2000 images and 233 JPEG images.). Experiments show that the proposed metrics match with Human Visual System better than SSIM and VIF in pixel domain.

Acknowledgement:
This paper was supported by the NSF grant # 0619069, the Development of an Interdisciplinary Arkansas Emulation Laboratory and by the System Research and Application (SRA) Contract # S280000131. The authors would like to thank the colleagues from QBASE in Dayton, Ohio for useful comments and suggestions.

References:

[1] A.Moorthy, A. Bovik, "Visual Importance Pooling for Image Quality Assessment", *IEEE Journal of Selected Topics in Signal Processing*, Vol 3, No 2 April, 2009.

[2] H.R. Wu, K.R. Rao, "Digital Video Image Quality and Perceptual Coding", *CRC Press*, 2006.

[3] L. Itti C. Koch and E. Niebur, "A model of saliency based visual attention fro rapid scene analysis", *IEEE Trans. Pattern Anal. Mach. Intell*, vol.20, no.11, pp.1254-1259, Nov 1998.

[4] U. Rajashekar , A. Bovik and L. Cormack, Caffe, "A gaze attentive fixation finding engine", *IEEE Trans. Image, Processing*, vol 17, no 4, pp 564-573, Apr 2008.

[5] H. R. Sheikh, M. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

[6] H.R. Sheikh.and A.C. Bovik, "Image information and visual quality," ***IEEE Transactions on Image Processing***, vol.15, no.2,pp. 430- 444, Feb. 2006.

[7] P.G Schyns and A. Oliva, "From blobs to boundary edges: evidence for time and spatial scale dependent scene recognition", *Psychol Sci*, 5, 195-200, 1994.

[8] A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope", *Int.J.Comut Vision 42*,145-175, 2001

[9] M.Milanova , S. Rubin, R. Kountchev, V. Todorov, R. Kountcheva, "Combined visual attention model for video sequences", *IEEE ICPR 2008* , 1-4.

[10] A. Bell and T. Sejnowski, "The 'independent components' of natural scenes are edge filters", *Visual Research*, 37 (3), 3327-3338, 1997.

[11] B. Olshausen, "Sparse Codes and Spikes" in *Probabilistic Models of Perception and Brain Function*, R.P.N. B. Rao, A. Olshausen, M. Lewicki, Eds, *MIT Press*, 2001.

[12] R. Achanta, S. Hemami, F. Estrada and S. Süsstrunk, "Frequency-tuned Salient Region Detection", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

[13] Qiang Li and Zhou Wang, "Video quality assessment by incorporating a motion perception model," in Image Processing, 2007. ICIP 2007. *IEEE International Conference on, 2007*, vol. 2, pp. 173–176.

[14] Qi Ma, Liming Zhang, "Image quality assessment with visual attention", *ICPR 2008*, 1-4.

[15] P. L. Callet, F. Autrusseau, 'Subjective quality assessment IRCCyN/IVC database', http://www.irccyn.ec-nantes.fr/ivcdb/.

[16] H.R. Sheikh, Z.Wang, L. Cormack and A.C. Bovik, "LIVE Image Quality

Assessment Database Release 2", http://live.ece.utexas.edu/research/quality.
Appendix 1-B: "Objective Video Quality Assessments for Tracking Moving Objects from Video Sequences."

"Objective Video Quality Assessments for Tracking Moving Objects from Video Sequences." Engin Mendi, Mariofanna Milanova, Yinle Zhou, John Talburt – University of Arkansas at Little Rock. Presented at the 9th WSEAS International Conference on Signal Processing, Robotics, and Automation; Cambridge, UK; February 20-22, 2010.

Abstract – *Video quality assessment has a great importance in several image processing applications. Recently, various objective video quality metrics have been proposed to predict the human visual perception and to achieve high correlation with the human perception of the image quality. In this paper, a novel objective quality metric is proposed for tracking moving objects in video sequences. The proposed metric particularly considers the moving object in video sequences as visually important content. Foreground masks are produced by background subtractions based on an approximate median filter. Existing metrics are then modified by the weighting factors of the foreground masks. Our results show that our metrics have better performance than existing objective metrics.*

Objective Video Quality Assessment for Tracking Moving Objects from Video Sequences

E. MENDI
Department of Applied Science,
University of Arkansas at Little Rock,
Little Rock, AR, 72204
USA
esmendi@ualr.edu

M. MILANOVA
Department of Computer Science,
University of Arkansas at Little Rock,
Little Rock, AR, 72204
USA
mgmilanova@ualr.edu

Y. ZHOU
Department of Information Science,
University of Arkansas at Little Rock,
Little Rock, AR, 72204
USA
yxzhou@ualr.edu

J. TALBURT
Department of Information Science,
University of Arkansas at Little Rock,
Little Rock, AR, 72204
USA
jrtalburt@ualr.edu

Abstract: Video quality assessment has a great importance in several image processing applications. Recently, various objective video quality metrics have been proposed in order to predict the human visual perception and to achieve high correlation with the human perception of the image quality. In this paper, a novel objective quality metric is proposed for tracking moving objects in video sequences. The proposed metric particularly considers the moving objects in video sequences as visually important content. Foreground masks are produced by background subtraction based an approximate median filter. Existing metrics are then modified by the weighting factors of the foreground masks. Our results show that our metrics have better performance than existing objective metrics.

Key-Words: Video Quality Assessment, Background Subtraction, Tracking Moving Objects from Video Sequences.

1 Introduction

Video quality assessment (VQA) is important study for many applications .The industry's need for accurate and consistent objective video metrics has become more critical with new digital video applications and services such as Internet video, surveillance, mobile broadcasting and Internet Protocol television (IPTV).

VQA methods fall into two categories: subjective assessment by humans and objective assessment by algorithms. Objective quality metrics are algorithms designed to characterize the quality of video and predict viewer opinion. Different types of objective metrics exist as illustrated in paper [1]. In the image processing community more than 50 years mean squared error (MSE) are being used as quasi – standard fidelity metrics. The MSE sill continues to be widely used as a signal fidelity measure, but at the same time there are recent studies that have developed more advanced signal fidelity measures, especially in applications where perceptual criteria might be relevant. How well an algorithm performs is defined by how well it correlates with the human perception of quality. It is interesting to demonstrate how the video quality is measured for video records where the task is tracking moving objects. It is intuitively obvious that we need to use weighting factors for different regions and measure video quality. Only a small number of existing VQA algorithms detect motion and use motion information directly [2]. A heuristic weighting model is combined with the structural similarity (SSIM) based quality assessment method. The authors use the fact that the accuracy of visual perception is significantly reduced when the speed of motion is extremely large. In [3] a set of heuristic fuzzy rules are proposed that use both absolute and relative motion information to describe visual attention and motion suppression. In [2] the authors use the fact that the human visual system (HVS) is an optimal information extractor. In recent years, it has become clear that many problems in perception organization are difficult to solve without introducing the contextual information. We see and hear the world in terms of meaningful causal interactions. Barlow's hypothesis is that the purpose of early visual processing is to transform the highly redundant sensory input into more efficient factorial code. A perceptual system should be organized to transmit maximum information. This hypothesis and the other hypothesis that humans use consecutive approximation with increasing resolution for the selected regions of interest are implemented in [4].

Inspired from new cognitive image representation framework [4] we have developed improved VQA algorithm incorporating the model of motion as spatiotemporal weighting factors. In our video quality measure the weight increases with the information content and decreases with the perceptual uncertainty. The rest of the paper is organized as follows: Section 2 provides brief information about the tracking algorithm. An overview of existing and proposed quality metrics are presented in Section 3. We present the results of our approach in Section 4. Finally, in Section 5 the conclusions of this paper are summarized.

2 Tracking Algorithm

We have tracked moving cars from our traffic video data using background subtraction based on approximate median filter. Since the background is more likely to appear in our traffic data, approximate median, which is computationally efficient and fast, can be used. In this approach, background pixel is incremented by 1, if the input pixel is greater than the corresponding background. Similarly, if the input pixel is smaller than the background pixel, then corresponding background pixel is decremented by 1. In this way, background pixels converge to a value, where half of the input pixels are greater and the half of them is smaller than this value, which is the median. [5]
Background B is estimated at a time t, for input frame I as follows:

$$B(x, y, t) = median\{I(x, y, t-i)\} \qquad (1)$$

where $i \in \{0,1,...,n-1\}$ and n denotes the previous frames.

Once background is estimated, foreground mask is obtained by applying a threshold τ to the absolute difference of estimated background and input frame:

$$|I(x, y, t) - median\{I(x, y, t-i\}| > \tau \qquad (2)$$

Estimated background and foreground mask of out traffic video data for $n = 20$ is given in Fig. 1.

3 Objective Video Quality Assessment

The most reliable way to measure of Video quality is perceptual quality based on subjective evaluation by orienting on human visual system (HVS). Subjective measures are determined by Mean Opinion Score (MOS) which relies on human perception. On the other hand, objective metrics are also very valuable to make meaningful quality evaluations. They are based on mathematical measurements which are practical to apply without need of human observers. Such methods are widely used in various image processing applications, including filter design, image compression, restoration, denoising, reconstruction, and classification [6]. Objective quality metrics can be classified into 3 metrics: Full Reference (FR), Reduced Reference (RR) and No Reference (NR). All these metrics are based on the availability of original non-distorted reference image which will be compared with the corresponding distorted image. In FR case, reference image information is available; in RR case, partial information of reference image is known and no information about the reference image is available in the NR case.

3.1 MSE

Consider two images $x = \{x_i \mid i = 1,2,..., N\}$ and $y = \{y_i \mid i = 1,2,..., N\}$ where N is the number of pixels and x_i and y_i are the i th pixels of the images of x and y, respectively; the MSE between these two images is:

Distribution A; Approved for Public Release; Distributed Unlimited; 88ABW/PA cleared 24 September 2012 as 88ABW-2012-5092.


(a) (b)

Fig. 1: a) Estimated background, b) foreground mask

$$\mathrm{MSE}(x,y)=\frac{1}{N}\sum_{i=1}^{N}(x_i-y_i)^2 \quad (3)$$

MSE is widely used as it is parameter free, computationally simple and mathematically convenient in the context of optimization. It also represents image energy measure that energy is preserved after any orthogonal linear transformation, such as the Fourier transform. However, MSE does not fit precisely with the perceived visual quality. Distorted images with the same MSE may have different visibility. [6] [7]

3.2 SSIM

To overcome limitations of MSE, a new objective quality metric SSIM [7] has been proposed. SSIM correlates well with human subjective perception [8]. Consider two images $x=\{x_i \mid i=1,2,...,N\}$ and $y=\{y_i \mid i=1,2,...,N\}$ where N is the number of pixels and x_i and y_i are the ith pixels of the images of x and y, respectively. SSIM-$\mathrm{SSIM}(x,y)$ combines three comparison components, namely luminance-$l(x,y)$, contrast-$c(x,y)$ and structure-$s(x,y)$:

$$\mathrm{SSIM}(x,y)=f(l(x,y),c(x,y),s(x,y)) \quad (4)$$

Luminance, contrast and structure comparisons are defined as follows:

$$l(x,y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1},\quad C_1=(K_1L)^2$$

(5)

$$c(x,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2},\quad C_2=(K_2L)^2$$

(6)

$$s(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3},\quad C_3=\frac{C_2}{2}$$

(7) where μ_x, μ_y, σ_x, σ_y and σ_{xy} are means of x and y, variances of x and y and correlation coefficient between x and y. K_1 and K_2 are scalar constants that $K_1, K_2 << 1$ and L is the dynamic range of the pixel values. Finally, SSIM index yields to:

$$\mathrm{SSIM}(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \quad (8)$$

3.3 Weighted Objective Quality Metric

In human visual system, the importance of a visual event should increase with the information content, and decrease with the perceptual uncertainty [2], we incorporated foreground mask (2) as weighting function into the MSE and SSIM metrics to measure the motion feature of the moving car. At a time MSE is $\mathrm{MSE}(x,y,t)$ and SSIM is $\mathrm{SSIM}(x,y,t)$. The weighting function is:

$$w(x,y,t)=\left|I(x,y,t)-median\{I(x,y,t-i\}\right|>\tau$$

(9)

We define weighted MSE as wMSE and weighted SSIM as wSSIM as follows:

$$w\mathrm{MSE}=\frac{\sum_x\sum_y\sum_t w(x,y,t)\mathrm{MSE}(x,y,t)}{\sum_x\sum_y\sum_t w(x,y,t)} \quad (10)$$

$$w\mathrm{SSIM}=\frac{\sum_x\sum_y\sum_t w(x,y,t)\mathrm{SSIM}(x,y,t)}{\sum_x\sum_y\sum_t w(x,y,t)}$$

(11)

4 Experimental Results

We demonstrated the weighted new objective quality metrics on an intuitive example. We used a traffic video data containing 23 frames from a ground sensor camera. We distorted the original reference video generated from 3 different processing: Blurring, Salt and Pepper noise and JPEG compression. Each process has also 3 distortion amount. Distortion types and amounts are summarized in Table 1.

A sample frame image from the video data and associated distortions are depicted in Fig. 2.

Fig. 3 shows the results of objective VQA. Fig.3.a, 3.c and 3.e show the MSE and wMSE scores of blurred, salt and pepper noise and JPEG compression distortions respectively. Corresponding SSIM and wSSIM scores are given in Fig.3.b, 3.d and 3.f. The x axis in the figures denotes the frame index (time), while the y axis denotes MSE & wMSE or SSIM & wSSIM. As shown in the figures, weighted metrics are more realistic and correlated with human perception. For instance, since there is no moving car in the first frame, MSE and SSIM give wrong scores, while weighted metrics give 0.0 and 1.0, respectively, as they give importance to only moving content. Similarly, in other frames, wMSE values are less than of MSE, and wSSIM values are greater than of SSIM. This is because visually important content such as the moving car is more considered by wMSE and wSSIM.

5 Conclusions

In this paper, we presented a novel objective quality assessment metric. In proposed metrics, moving objects from video sequences are particularly considered as visually important content. Background subtraction based on approximate median filter is used for tracking the moving objects. Then foreground masks are computed from the absolute difference of estimated background and input frame. Existing metrics MSE and SSIM are modified by the weighting factors of the foreground masks. We applied our approach to a traffic video data from a ground sensor. Our results show that our metrics are more realistic and correlated than existing metrics. In the future we will develop a subjective quality assessment to validate our metrics with human subjective perception.

Acknowledgement:

This paper was supported by the NSF grant # 0619069, the Development of an Interdisciplinary Arkansas Emulation Laboratory and by the System Research and Application (SRA) Contract # S280000131. The authors would like to thank the colleagues from QBASE in Dayton, Ohio for the reference video sequences provided.

References:

[1] S. Winkler, "Video Quality Measurement Standards – Current Status and Trends", *ICICS*, 2009.

[2] Q. Li, Z. Wang, "Video quality assessment by incorporating a motion perception model", *ICIP'2007*, pp 173-176.

[3] Z.K. Lu,W.Lin,X.K. Yang, E.P,Ong and S.S. Yao, "Modeling visual attention's modularity aftereffects on visual sensitivity and quality evaluation", *IEEE Trans. Image Processing*, vol 14, pp 1928-1942, 2005.

[4] R. Kountchev , S. Rubin , M. Milanova, V. Todorov , R. Kountcheva, "Cognitive image representation based on spectrum pyramid decomposition Source", *WSEAS International Conference on Mathematical Methods and Computational Techniques in Electrical Engineering* , 2008 , pp230-235.

[5] S.-C. Cheung and C. Kamath, "Robust techniques for background subtraction in urban traffic video," *in Proc. Video Communications and Image Processing, SPIE Electronic Imaging*, San Jose, CA, USA, January 2004.

[6] Z. Wang and A. C. Bovik, "Mean Squared Error: Love It or Leave It?", IEEE *Signal Processing Magazine,* January 2009.

[7] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", *IEEE Transactions on Image Processing,* vol.13, no.4pp. 600- 612, April 2004.

[8] H. R. Sheikh, M. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process*., vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

Distortion Type	Distortion 1	Distortion 2	Distortion 3
Blurring	fil. size=6, std. dev = 6	fil. size=8, std. dev = 8	fil. size=10, std. dev = 10
Salt and Pepper	d (noise density) = 0.01	d (noise density) = 0.03	d (noise density) = 0.05
JPEG Compression	compression = 50%	compression = 70%	compression = 90%

Table 1: Distortion processing and amounts

Fig. 2: a) Sample reference frame, b) blurred of size 10 with standard deviation 10, c) salt and pepper noise with noise of 0.05, d) JPEG compression with 90%

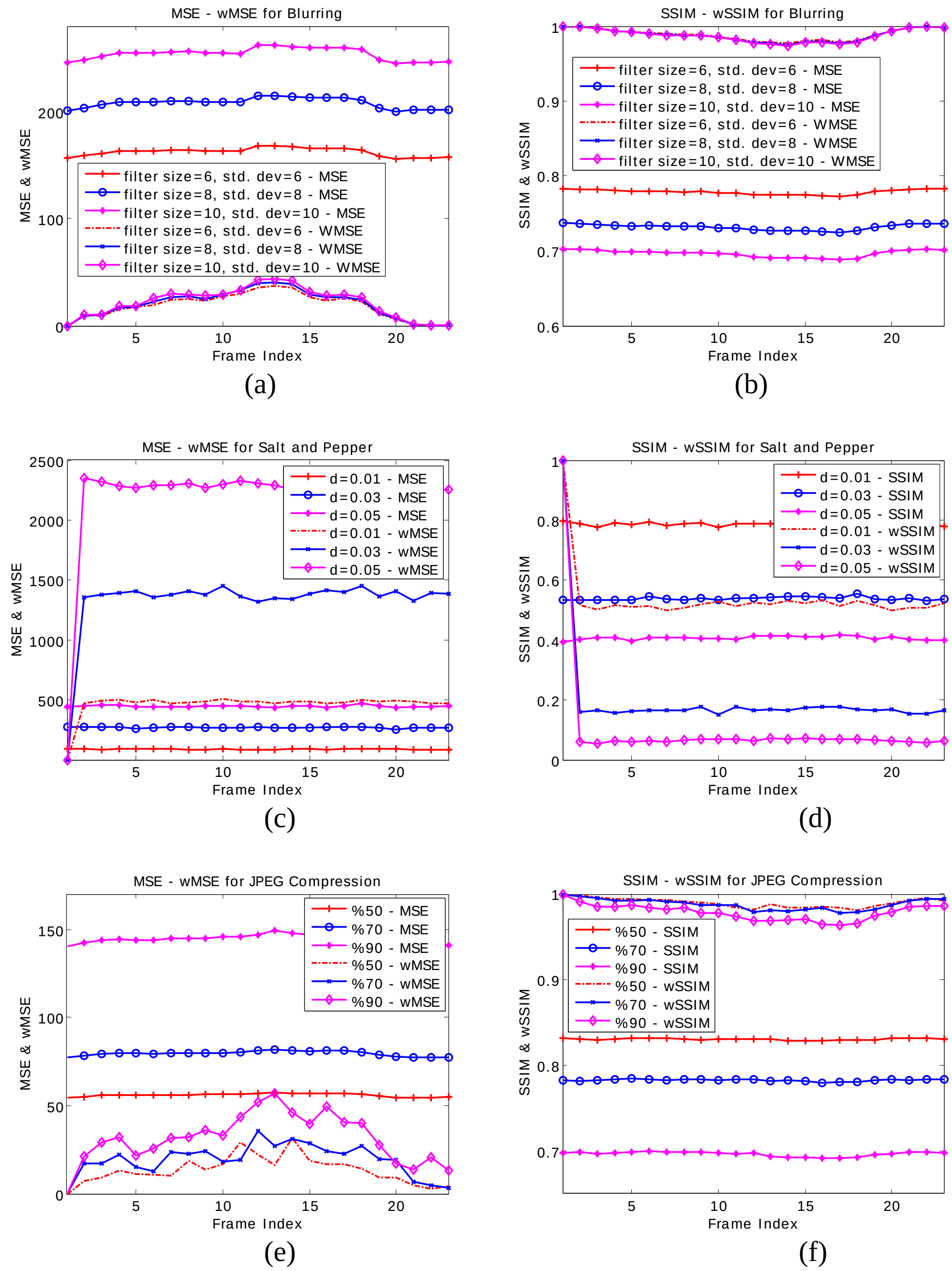


Fig. 3: Objective VQA plots on a test video containing 23 frames, a) MSE and MSE with proposed weighting method for blurring distortion, b) SSIM and SSIM with proposed weighting method for blurring distortion, c) MSE and MSE with proposed weighting method for salt & pepper effect, d) SSIM and SSIM with proposed weighting method for salt & pepper effect, e) MSE and MSE with proposed weighting method for JPEG compression, f) SSIM and SSIM with proposed weighting method for JPEG compression

APPENDIX 2: TRACK 2 – PROTOTYPE THE UTILIZATION OF INTERACTIVE 3-D INFORMATION VISUALIZATION IN THE LAYERED SENSOR DOMAIN

Appendix 2-A: "Mapping Realities: The Co-Visualization of Geographic and Non-Spatial Textual Information"

"Mapping Realities: The Co-Visualization of Geographic and Non-Spatial Textual Information." O. Isaac Osesina, M. Edward Tudoreanu, Cecilia Bartley – University of Arkansas at Little Rock. Presented at the 2010 International Conference on Modeling, Simulation, and Visualization Methods; Las Vegas, Nevada; July12-15, 2010.

Abstract – *This paper presents an approach for visualizing unstructured text via a geospatial milieu. The logical associations between textual information and geospatial data are used to determine geographical placement of keywords from the text. Interaction of the user with the additional information category is the geographical information system (GIS) application does not require additional effort other than the traditional zooming and panning on a map, thus making the non-spatial text a seamless component of GIS. The spatial placement of tweets keywords exposes potential relationships between tweets and geographic area that otherwise might not be visible. Out contribution resides in techniques for extracting geospatial data, in assigning location to non-spatial text based on the logical association to geo-located data and in designing visualization techniques for conveying the textual information at various levels of detail.*

Mapping Realities: The Co-Visualization of Geographic and Non-spatial Textual Information

O. Isaac Osesina, M. Eduard Tudoreanu, and Cecilia Bartley
Department of Information Science, University of Arkansas at Little Rock, Little Rock, AR, USA

Abstract – *This paper presents an approach for visualizing unstructured text via a geospatial milieu. The logical associations between textual information and geospatial data are used to determine geographical placement of keywords from the text. Interaction of the user with the additional information category in the geographical information system (GIS) application does not require additional effort other than the traditional zooming and panning on a map, thus making the non-spatial text a seamless component of GIS. The spatial placement of tweets keywords exposes potential relationships between tweets and geographical areas that otherwise might not be visible. Our contribution resides in techniques for extracting geospatial data, in assigning location to non-spatial text based on the logical associations to geo-located data, and in designing visualizations techniques for conveying the textual information at various levels of detail.*

Keywords: Twitter Visualization, GIS, Situational Awareness (SA)

1. Introduction

The advancements in information technology among other things have greatly increased the amount and type of information available to individuals and organizations. Due to the diverse source and scope of the available information theoretically, complete understanding of the world could be achieved by combining these diverse information sources. This provides an opportunity for applications that are able to utilize information from multiple sources to create an enriched user experience. Map applications powered by tools like Web 2.0 and geographical information systems (GIS) combine traditional spatial information such as country, city, street, river, and topography together with information such as organization locations, events, and social networking activities in order to create an enriched visual experience for users.

Apart from the drawback that the available information exists in different formats, joining various pieces can be very complicated, especially when data is primarily created for different purposes. Consider the task of presenting both a news article about an ongoing pandemic in City A and census data about the same city. Moreover, given that up to 80 percent of valuable information has been estimated to be in unstructured form [1] [2] and that most applications require strict data format; utilizing a majority of the existing information can be very challenging. Furthermore, the integration of information sources with different data quality requirements may leave the final information product vulnerable to complex quality problems.

This paper presents an approach to integrating multiple, heterogeneous data sources and delivering the final information product in an easily navigable and comprehensible format. We devised a technique of simultaneously processing and analyzing textual and map information in order to create a common visualization. The interaction requirements of our techniques are virtually the same as typical GIS exploration. Furthermore, we examine an information visualization method that helps users easily digest the integrated information

Our approach exploits the logical linkage between the non-spatial textual data and the GIS information. The exact origin on Earth of the text may never be known, but the meaning refers to specific places. This paper describes techniques both for discovering the linkage and for visualizing the distribution of important textual keywords on the map. We believe that for most tasks and applications, the logical relationship between GIS and text will prove important, especially because it may be the only type of relationship available to infer.

A proof of concept system that introduces social networking site information to a GIS system was developed with the goal of providing usable data to first responders, regional administrators, marketers etc. In particular, we combined information from Twitter, one of the fastest growing online social networking services used to broadcast short

Distribution A; Approved for Public Release; Distributed Unlimited; 88ABW/PA cleared 24 September 2012 as 88ABW-2012-5092.

messages (a.k.a. tweets), and World Wind, a GIS-based, Google Earth-like application developed by NASA. The information from Twitter was strategically injected into World Wind in such a way that the information can be easily correlated with geographical space in order to convey the spatiality of the buzz on Twitter. Both main sources of information were processed and analyzed simultaneously in near real-time and presented to users in an interactive and easily understandable format in such a way that the user is not required to perform any additional activity other than the usual pan, zoom, and hovering needed for the operation of a GIS application.

Section 2 contains the background and related work relevant to this research. In Section 3, we describe our design of joining Twitter and World Wind information, automatically creating queries to extract information from Twitter, the analyses of the textual data, and displaying the information on the map. Section 4 details a user perspective on the opportunities and drawbacks of integrating these two forms of information. The paper ends with the Conclusions and Future Work in Section 5.

2. Background

This section describes related work, which falls within three broad categories: data fusion, also known as "metasearch", geo-tagging of various information into a GIS environment, and geo-visualization of Twitter data. We also cover the tools employed in our research: NASA World Wind, the GIS environment, and social networking sites, such as Twitter.

Data fusion strategies are outlined by Dong and Naumann [3], who classify various strategies of integration. They are generally based on the existence of meta-data, and do not focus on spatial or visual techniques. Wu and Cretani [4] adaptive approach to weighted data fusion introduces new algorithms for the input systems involved in the fusion process. Shou and Sanderson [5] focus on search engines and present an approach to cross rank data from multiple engines. The entities involved in the fusion are in fact the same, which is not the case in our work with GIS and non-spatial text. The information retrieval problem is looked at from another viewpoint by Efron [6] whose probabilistic framework assumes the level of interest declines with the quantity of information. This study focuses on alleviating that problem by combining information with very different characteristics, capable of presenting complementary views of the world

Geo-tagging is employed by tools such as MapCruncher [7], World Explorer [8], and MetaCarta [9] to allow people to include textual, pictorial, or other types of tags on a GIS data set. MapCruncher [7], designed by Microsoft, attempts to offer user flexibility in visualization by placing overlays on top of standard maps. These layers allow the user to drill down to a more granular view of geospatial information, but the basic premise is that the overlays are manually geo-located by a user. World Explorer's [8] use of textual features from geo-tagged Flickr data creates a dynamic map based on geo-referenced coordinates. The accuracy of the map relies on the users input of correctly tagged images, and cannot handle non-tagged data. MetaCarta's [9] products allow users to design mashups based on their own content. While the approach allows the inclusion of textual information, the visualization of that information lacks the ability to offer an overall view of the data. Users are left to click on each tag of interest in order to read the information. Our approach intrinsically offers an overall view of the Twitter buzz around the country or the world.

Geographic information systems (GIS) are used to manage and present location based data. In contrast to the static maps, GIS allows users to interactively communicate, analyze, and edit the data. Due to this dynamic nature of the GIS, several applications that use geographically referenced data has been improved or developed e.g., GPS, remote sensing, and aerial photography. World Wind is a free, open source virtual globe Java application developed by NASA. It allows users to remotely access NASA, USGS, and publicly available GIS data such as satellite imagery, aerial photography, topographic maps, road maps and political boundaries which can each be viewed as different layers on the map [10]. It was the GIS application of choice in our application due to its open source nature and the relative ease of creating a new custom layer for displaying our information.

Interactive visualizations of Twitter come in two flavors, geospatial and abstract, not based on a map. Just Landed [11] extracts landing locations from tweets that contain the phrase "just landed in…" and displays those locations on a map. However, Just Landed [11] does not display any textual information from tweets. TrendsMap [12] relies on the user's profile in Twitter, to display keywords, but in the event that the user is not "home" the information is incorrect. Twitter's universe can also be visualized by Monitter [13], Tweet Wheel [14], and Twitter StreamGraphs [15], are all non-geo-coordinated visualizations. While these tools are unique in their approach to visualization, they do not display spatial distribution on a map nor can they convey situational awareness for various parts of the country/world.

We chose Twitter as the social networking site for our prototype because of its microblogging feature. It allows users to send messages among a group of "follower" or broadcast short messages (a.k.a. tweets) which can be accessed via different media e.g., internet, phone and custom applications. Since its creation in 2006, it has become one of the largest social networking services. It grew at the rate of 752% to a total of 4.43 million unique visitors in 2008 [16] and more than 75 million user in January 2010 [17].

Our prototype leverages on Twitter, as well as, on GIS data to provide situational awareness to a wide variety of audiences from first responders to marketers.

3. Approach Overview

This section describes our approach to assigning spatial positions for non-spatial data, in particular unstructured textual data. In this section, we describe the reasoning behind our choice of linkage between the two categories of information, extraction of relevant information from both sources and the placement of textual data as a word cloud visualization on the map.

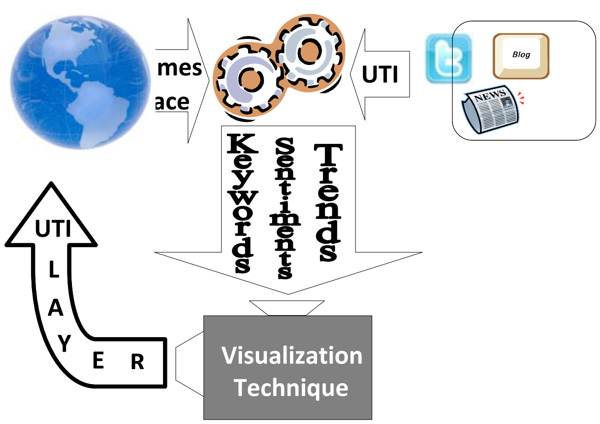


Figure 1: GIS information is extracted and used to assign location to unstructured text. Additional processing takes place to determine important keywords, sentiments and trends which are then included in a visualization to be displayed on the map.

3.1 Logical linking of data sets

By nature, unstructured text does not have standard geographical coordinates. Hence, the geospatial representation of textual information is not typically a straight forward endeavor. Therefore, a critical task in integrating textual information with geospatial information is determining the best algorithm for joining different categories of data. The algorithm depends on factors such as the intended use of the final information product, the scope of the textual information, and the available data manipulation technologies. For example, to develop an application that uses the IEEE publications database to determine the concentration of information visualization researchers in the US. The name of organizations to which artificial intelligence researchers are affiliated would be a good attribute to join research publications and GIS information. Generally speaking, the question may be more complicated in the absence of a clear structure of the document, and a lack of a spatial attribute. For such unstructured data, a more logical or semantic search can be performed to determine the textual data for which geo-coordinates can be inferred.

Our proof of concept is focused on geospatially representing Twitter information on World Wind in an application that can be used for example by first responders to increase their awareness of a theatre of operation. There is no direct widely available attribute that can be used to link the two categories of data. Although the geo-location feature of tweets (contains the geo-coordinates of the tweet origin) can be used to join both information categories, there are a couple of quickly visible drawbacks to this approach.

Firstly, only a fraction of tweets have values for this feature (according to eWeek.com [18] only 0.23%), which would render the vast majority of Twitter data un-joinable. The second drawback is more specific to the intended use of the application. For the described usage of our application, the information contained in the tweet body provides more information about a certain location than the origin of the tweet. For instance, a tweet might be sent from a hotel room at point A, but it concerns events happening in the home city of B. Hence, the logical link between GIS and information contained in the tweet better serves the intended purpose of our application than a join based on the tweet origin.

Considering that our main aim for introducing Twitter information into World Wind is to get a feel for the buzz about geographical locations contained in the Twitter chatter, we employed place/location names as the basis for exploring the logical linkage. This approach gives us the ability to assign to tweets geographical coordinates of location(s) whose name they contain. In general, any location for which geo-coordinates exist can be used (e.g., state, counties, cities, landmarks, organizations and street) as can databases with various public servants names such as mayors or governors. We used the location names available to the PlaceName layer of World Wind. The layer contains the names of continents, countries, towns, and cities as well as their geo-coordinates.

3.2 *Extracting spatial attributes/information of interest*

This section describes the process of extraction of the required information from both the GIS and textual sources. Systematically extracting and requiring information not only makes the application efficient, but also makes it user friendly.

Given the large geographical distances that can be covered in a relatively short time on a GIS application like World Wind, a lot of place names must be extracted from the PlaceName layer and used as a query argument to extract tweets of interest from Twitter. We approached this task by using some of the existing mechanisms in World Wind to decide what information to request/extract from its databases. In particular, we determined from the World Wind interface the geographical area in view of the user. Various place names are associated with that area at different levels of details as shown in Figure 2.

Our approach is to extract names from multiple levels of detail, even if World Wind does not currently render that level (because the user may be at a high altitude). The place names collected from World Wind are used in the Twitter query to obtain the tweets that contain those names. Effectively, the user queries Twitter simply by flying from one place to the other on the globe. Consequently, the number of places flown over is directly proportional to the number of queries performed on Twitter. This approach has the general effect that the user can explore multiple information sources while expending only the same usual amount of effort required for maneuvering on World Wind.

Although World Wind, like many other GIS applications, manages the level of details displayed for a geographical area with respect to altitude (altitude of user's view), this level of detail depends on several factors that might not contain enough information for extracting data from the unstructured textual information (UTI) source. For example, if the GIS shows information at the country level but the UTI does not contain country names, then it would be impossible to extract information from the UTI based on that particular string (country name). We used information with finer details beyond that displayed on the map for querying Twitter. For example, if the user is at the country level, we dig deeper to obtain further information about the states, or cities in view of the map.

This level of detail allows us to extract more data from Twitter (and presumably more information) about the location. After this information is analyzed, the results are aggregated with respect to the geographical space and presented to the user. This also has the advantage that information from Twitter would have already been requested (and possibly obtained and analyzed) before the user reaches lower altitudes.

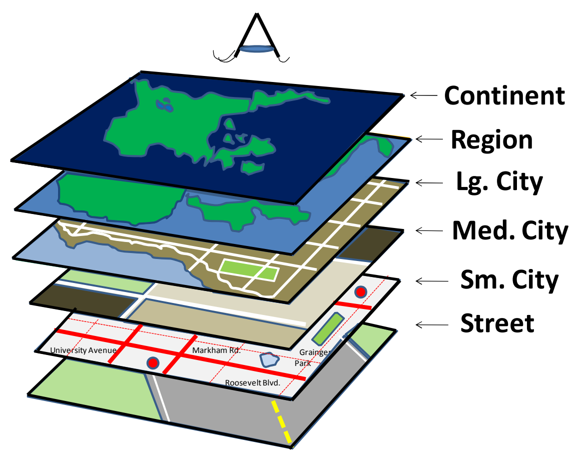


Figure 2: Information layers according to level of details. The information displayed to a viewer may be less detailed than the information used to query a textual data source such as Twitter.

In order for the user to access the knowledge contained in the UTI without having to read through several lines of text (especially when s/he is pressed for time) the UTI should be analyzed and presented in an easily interpretable format to the user. The type(s) of analyses to be carried out depends on the intended use of the final product, the type of UTI, as well as the technology. We performed three key analyses on the set of tweets obtained from each query result namely: keywords analysis, sentiment analysis, and trend analysis

We used the keywords analysis to get an overview of the main topics of the discourse in the query results for a particular place. The analysis was performed by identifying the most frequent words other than English language stop words[1] and the query argument[2] contained in the tweets. The result of the analysis contains each word in the result set frequency. We attempted to use the sentiment analysis to capture the emotion/mood (happy, sad, and panic) expressed in the tweets. Depending on the use of this application, the mood about a place can trigger different actions. For example, detection of a panic emotion may prompt the police to increase physical presence. We used a pre-compiled list of words that signal each of the three moods that we analyzed. The mood of the query result is decided by determining which of the buckets of pre-compiled emotion words is most represented in the entire keywords set (i.e., the bucket with the most number of words that can be found in the keywords set). The trend analysis is simply a way to create some persistence in the data analysis. It depicts fluctuations (if any) in the Twitter mood and most frequent keywords in a place over time. This is important as it can help in identifying unusual patterns in the information obtained for a place.

3.3 Display on the map

Two main types of strategies for rendering keywords in a spatial environment are presented, and they differ in whether single points or entire areas are considered as anchors for displaying Twitter data. The visualization strategies balance the two competing goals of presenting a large number of Twitter keywords and of creating a clear, easy to understand view.

Point-based strategies start by directly assigning the location of each query (e.g., city, street name) to the Twitter keywords that correspond to that query. This simple approach results in keywords being displayed on top of each other and, if visible, on the query name (location). This problem can be corrected by artificially spreading out the keywords and even by changing their font size as needed. Either random alteration of the keyword placement or some regular, geometric pattern can be implemented.

Point-based strategies display very precisely the association of queries to keywords, but aggregation occurs through spatial placement, and may hide some important keywords in favor of relatively unimportant ones. This "visual aggregation" occurs on the map in the sense that for a given area, such as a state, all keywords about the cities in that state are shown next to each other. A lack of explicit aggregation may lead to a situation in which keywords with low frequency may be displayed while some relatively frequent keywords are not visible. Consider the case of two neighboring cities, one small and the other large. The small city may be able to display all its associated keywords, even those that only occur once or twice in tweets. The large city may have a large number of keywords all with a high frequency, yet due to space constraints, only a fraction of those keywords are shown. This is exactly the case of unimportant keywords being visible (around the small city) while frequent terms are cut out (around the large city).

The second type of strategy, area-based, can emphasize overall aggregation of keywords. All the Twitter terms that fall within a given area are considered together, and only the most important ones are displayed. The size of the area is dependent on how far the user is from Earth's surface, and we implemented this using standard GIS tiles. Keywords are placed around the area in such a way to avoid overlap, and they are regarded as part of the area rather than belonging to a point on the map. This provides a larger real-estate for placement, and can alleviate repetition of keywords.

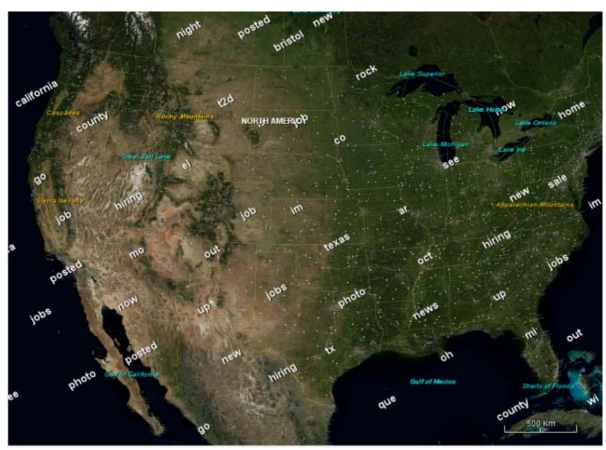


Figure 3: Twitter keywords placement approximation at the country level view on World Wind.

There are three types of techniques for spreading out Twitter terms in an area. The first technique relies on random placement; the second

[1] The language of the stop words is dependent on the language of the text.
[2] The query argument (place name) was not included in the analysis because, it does not provide any additional information and it is contained in the all tweets in the query result.

technique uses a geometric pattern, such as concentric circles (see Figure 4) or grids; and the third technique uses weighted-average placement of keywords. The first two techniques are computationally inexpensive, with the second being able to also convey the relative importance of the terms by, for example, placing the most important term in the center and using a pre-defined ordered placement after that. The third technique requires the use of either forces or virtual "bungee cords" to pull a keyword in its final position. The anchor points for the forces or cords are the location of the queries associated with the keyword (e.g., the cities in which the keyword is tweeted). This approach is similar to MonkEllipse [19], and it will result in popular/widespread keywords appearing in the center of the area because they are "pulled" in multiple directions towards most places in that area. The averaged-position may also lead to overlap, and requires an extra overlap-reducing step in which keywords lying on top of each other are spread around.

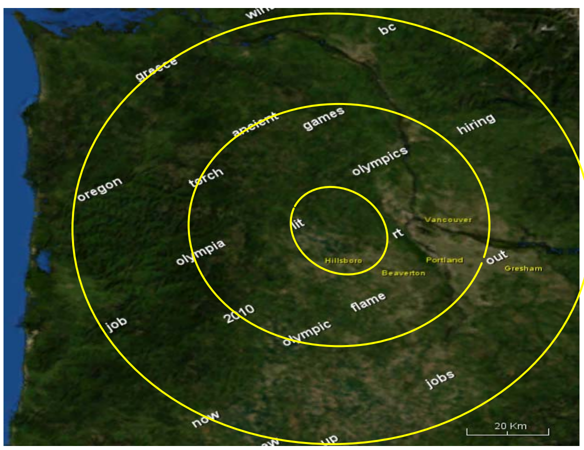


Figure 4: Example of a word cloud over Vancouver. The yellow circles are for illustration purposes only and do not appear in the visualization.

4. Opportunities and Challenges

A formal study of the benefits and limitations of our research is currently under way. During the design and implementation of the study, however, we were able to detect high level opportunities and challenges for the visualizations of GIS and Twitter data. The system was also demonstrated for several hours at the Air Force Research Laboratory at Tec^Edge [20] "Summer on the Edge", 2009.

One of the advantages of our approach is that the situation about different locations can be monitored in near-real time using keywords from analyzed live tweets feed. Moreover, the geographical information (map) help users to better interpret tweet data and estimate the level of believability. Finally, the automatic querying of Twitter, analyses of tweets and the aggregation of analyzed data enabled the users to obtain a lot of information without altering their usual activities of operating World Wind. The disadvantages include a high level of noise in Twitter information (e.g., excessive airport weather information), rate-limited querying of real-time tweets, and space constraints in displaying keywords.

An example of tweets as a source of information on current/real-time events was during the 2009 Iranian election crises [21]. We observed that some information obtained from tweets shown on the map became front page news the following day or was recently in the news (e.g., Pittsburg shooting and sweat lodge deaths). Twitter also provides a broad perspective of events concerning a community even if those events do not necessarily make it into the national media. The aggregation and analysis of this data in relation to geographical location can give a sense of the community discourse (buzz) which is valuable for situational awareness.

Although the integrity of a tweet source might be unknown, analyzing tweets in the context of location may help to discern the believability of its content. An example could be the keyword "drought" showing up in the Seattle area, which is typically quite rainy.

The use of keywords, charts and other visualization methods enables users to more easily acquire knowledge without having to read thousands of tweets within a relatively short time. This technique can be extended to other types of small text-based sources such as emails, news articles and even web search results.

Twitter information is inherently unreliable and error-prone because it originates from multiple people. According to Tom Anderson, a social media market researcher, Twitter can be described as a "Babylon of Spam" [22]. In order to reduce the "noise" level in our analysis, we filtered out certain tweets, such as automated weather reports. Furthermore, due to the relatively small number of characters allowed in a tweet – 120 characters [23] and the wide use of abbreviations and slangs, the analysis of tweets can be complex compared to other forms of UTI; for example, a traditional news article.

Twitter Inc. actively enforces a rate-limiting policy in which not more than 1,500 tweets can be returned by a query [24]. Furthermore, there is a limit of 250 API requests per hour (without special permission by Twitter) [25]. These restrictions on the amount of accessible information affect the timeliness of the information and possibly the quality of the analysis. This limitation can be overcome by locally archiving tweets. Another challenge is the limited real estate (geographical space) available for displaying the keywords in addition to other GIS information. In order to manage the available real estate, the number of keywords and/or size of its font needed to be reduced. Also, our implementation of an area-based placement of keywords may locate keywords in unrealistic areas on the map, especially at high altitudes such as in Figure 3. We found that 20 keywords per tile do not clutter the map (see Figure 4).

5. Conclusions and Future Work

We designed and implemented a technique for logically linking non-spatial textual data from Twitter with geo-located information. The linkage is based on extracting names of various GIS features, such as countries, states, or cities, and on querying Twitter for messages that refer to that information. The geographical coordinates of the place names used in the query are extended onto the tweets contained in the corresponding query results. The results are then analyzed for keywords (the buzz words), sentiment (mood of the tweeter's), and trends (changes in the patterns of keywords and sentiments over time). Finally, we detailed a number of visualization techniques for producing word clouds capable of presenting the most important keywords from Twitter.

Some of the opportunities of this research include the use of tweets as real time news, the high number of active users of Twitter, and the free availability of a substantial amount of social networking data. The challenges include the unknown integrity of the information contained in tweets, and the limited real estate available for displaying keywords.

We are in the process of conducting a more extensive study of Twitter on World Wind visualization. The results of the study would provide a more in-depth understanding of the advantages and limitations of this approach.

6. Acknowledgments

The research for this paper was supported in part by the US Air Force Research Laboratory Sensor and Human Factors Directorates at Wright Patterson Airbase (SRA Contract #S280000131). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of any of the above organizations or any person connected with them.

7. References

[1] Christopher C. Shilakes and Julie Tylman. "*Server and Enterprise Software - Indepth Report.*"; s.l. : Merrill Lynch & Co. Enterprise Information Portals. [ed.], November 16, 1998.

[2] Greg Goth. *"News: A Structure for Unstructured Data Search.";* IEEE Distributed Systems Online, Vol. 8, (1-4), January 1,2007.

[3] X. L. Dong, and F. Naumann. *"Data Fusion - Resolving Data Conflicts for Integration.";* ACM, Proceedings of the VLDB Endowment, Lyon, (1654 - 1655), 2009.

[4] S. Wu, and F. Crestani. *"Data Fusion with Estimated Weights.";* ACM, Proceedings of the Eleventh International Conference on Information and Knowledge Management, McLean, (648 - 651), 2002.

[5] X. M. Shou, and M. Sanderson. *"Experiments on Data Fusion Using Headline Information."* ACM, Proceedings of the Twenty-fifth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Tampere, (413 - 414), 2002.

[6] M. Efron. *"Generative Model-Based MetaSearch for Data Fusion in Information Retrieval."* ACM, Proceedings of the Ninth ACM/IEEE-CS Joint Conference on Digital Libraries, Austin, (153 - 162), 2009.

[7] J. Elson, J. Howell, and J. R. Douceur. *"MapCruncher: Integrating the World's Geographic Information."* SIGOPS Operating Systems Review, (50 - 59), 2007.

[8] S. Ahern, et al. *"World Explorer: Visualizing Aggregate Data from Unstructured Text in Geo-Referenced Collections."* ACM, Proceedings of the Seventh ACM/IEEE-CS Joint Conference on Digital Libraries, Vancouver, (1 - 10), 2007.

[9] J. Frank. *MetaCarta Inc. Web site.* [Online] January 1, 2001. [Cited: 02 01, 2010.] http://metacarta.com/. About Us: MetaCarta Inc.

[10] NASA. World Wind. *World Wind.* [Online] [Cited: February 9, 2010.]

http://ti.arc.nasa.gov/projects/worldwind/index.php.

[11] Jer Thorp, blprnt.blg. *Just-landed: Processing, Twitter, MetaCarta & Hidden Data.* [Online] May 11, 2009. [Cited: March 15, 2010.] http://blog.blprnt.com/blog/blprnt/just-landed-processing-twitter-metacarta-hidden-data.

[12] Stateless Systems. Trendsmap. *Trendsmap Real-Time Local Twitter Trends.* [Online] [Cited: March 15, 2010.] http://trendsmap.com/.

[13] Alex Holt. Monitter. *Soyrex .* [Online] [Cited: March 15, 2010.] http://monitter.com/.

[14] Manuel Lima. Social Networks - TweetWheel. *Visual Complexity.* [Online] [Cited: March 15, 2010.] http://www.visualcomplexity.com/vc/project.cfm?id=587.

[15] Max Kiesler. Twitter StreamGraph. *Neoformix .* [Online] [Cited: March 15, 2010.] http://www.maxkiesler.com/2009/07/20/twitter-streamgraph-visualization/.

[16] Adam Ostrow. Twitter's Massive 2008: 752 Percent Growth. *Mashable The Social Media Guide.* [Online] [Cited: February 9, 2010.] http://mashable.com/2009/01/09/twitter-growth-2008/.

[17] Robert J. Moore. New Data on Twitter's Users and Engagement. *The Metric System.* [Online] January 26, 2010. [Cited: February 26, 2010.] http://themetricsystem.rjmetrics.com/2010/01/26/new-data-on-twitters-users-and-engagement/.

[18] Clint Boulton. Twitter Use Tapering Off in the U.S., but Rising Overseas - Web Services, Web 2.0 & SOA. *eWeek.com.* [Online] January 15, 2010. [Cited: February 28, 2010.] http://www.eweek.com/c/a/Web-Services-Web-20-and-SOA/Twitter-Use-Declining-in-the-US-But-Rising-Overseas-592323/.

[19] Tzu-Wei Hsu, et al. "*MonkEllipse: Visualizing the History of Information Visualization.*" s.l.: IEEE Symposium on Information Visualization, (r9), 2004. 10.1109/INFVIS.2004.48.

[20] WBI-ICC (Tech^Edge). *Wright Brothers Institute - Innovation and Collaboration center.* [Online] [Cited: March 1, 2010.] http://www.wbi-icc.com/news/.

[21] Evgeny Morozov. Iran Elections: A Twitter Revolution? *The Washington Post.* [Online] June 17, 2009. [Cited: February 9, 2010.] http://www.washingtonpost.com/wp-dyn/content/discussion/2009/06/17/DI2009061702232.html.

[22] Tom H. C. Anderson. Reward for Being the Top Market Researcher on Twitter. *Tom H. C. Anderson - Next Gen Market Research.* [Online] Next Gen Market Researcher, February 3, 2010 . [Cited: February 4, 2010.] http://www.tomhcanderson.com/2010/02/03/reward-of-being-the-top-market-researcher-on-twitter/.

[23] Twitter Inc. Twitter Support: Frequently Asked Questions. *Twitter.* [Online] November 4, 2008. [Cited: February 29, 2010.] http://help.twitter.com/forums/10711/entries/13920-frequently-asked-questions.

[24] Twitter Search API. *Twitter.* [Online] [Cited: February 28, 2010.] http://search.twitter.com/api/.

[25] Twitter. Update and API Limits.*Twitter.* [Online] November 30,2008. [Cited: February 29, 2010.] http://twitter.zendesk.com/forums/10711/entries/15364.

[26] D. M. Boyd, and N. B. Ellison. *"Social network sites: Definition, history, and scholarship.";* Journal of Computer-Mediated Communication, Vol. 13, 2007.

[27] Neilsen. "*Global Faces and Networked Places-A Nielsen report on Social Networking's New Global Footprint.*" s.l.: The Nielsen Company, 2009.

Appendix 2-B: "Improving Information Quality of Textual Data by Geographical Reference."

"Improving Information Quality of Textual Data by Geographical Reference." O. Isaac Osesina, Cecilia Bartley, M. Edward Tudoreanu – University of Arkansas at Little Rock. Presented at the 2010 ALAR Conference on Applied Research in Information Technology; Conway, Arkansas; 9 April 9, 2010.

Abstract - *This paper provides an analysis of the information quality improvements that can be achieved by integrating multiple data sources with significantly dfferent characteristics. Multiple quality dimensions are covered in a case study in which geographically referenced Twitter information is introduced into a geo-spatial environment similar to Google Earth. The information was seamlessly integrated into the environment in such a manner that no additional activity is required from the user to explore Twitter data. We determined that overall value of both sets, but in particular Twitter, can be increased.*

Improving Information Quality of Textual Data by Geographical Reference

O. Isaac Osesina, Cecilia Bartley, M. Eduard Tudoreanu
Department of Information Science
University of Arkansas at Little Rock, Little Rock, AR 72204
{oiosesina, cxbartley, metudoreanu}@ualr.edu
February 22, 2010

Abstract

This paper provides an analysis of the information quality improvements that can be achieved by integrating multiple data sources with significantly different characteristics. Multiple quality dimensions are covered in a case study in which geographically referenced Twitter information is introduced into a geo-spatial environment similar to Google Earth. The information was seamlessly integrated into the environment in such a manner that no additional activity is required from the user to explore Twitter data. We determined that overall value of both sets, but in particular of Twitter can be increased.

Keywords**:** Information Quality, Visualization, Situational Awareness, Twitter, World Wind.

1. Introduction

The amount of information available to organizations and individuals has increased dramatically over the years [1]. This information contains data streams from sources as heterogeneous as text, images, or geographical information systems (GIS), .each potentially providing different perspectives into our changing world. These different, perspectives may need to be put together in order to create a better and more complete understanding of the world. The net worth of individual information sources can be improved by leveraging on each other's strengths and reducing each other's weaknesses, which is in line with Talburt's perspective on employing information quality (IQ) to help organizations maximize the value of their information assets [2].

Challenges to taking full advantage of this continuously-generated data include not only a limited amount of time and other resources available to process and digest this information (addressed in [3]), but also an apparent incompatibility between various data sources. One of the most notable problems is linking the ocean of textual information made available through various Internet channels to the geo-referenced data stemming from relatively verifable sources such as government statistics and sattelite imagery. The difficulty of correlating various sources translates into limited opportunities for data quality improvements.

This paper addresses the information quality improvements that are achieved by combining multiple, heterogeneous data sources. Our research explores a means by which multiple types of information can be processed and analyzed simultaneously. Furthermore, we examine the effect of integrating these multiple sources of information with different formats, integrity, and timeliness on data quality attributes of individual data.

We determined that the overall quality of "less reliable" information can be increased when put in the "context" of "more reliable" information. Note that IQ is a multidimensional space, and while a type of data may be better than another type in one dimension, it may also lag on other dimensions. The dimensions discussed here include Believability, Value-Added, Accuracy, Interpretability, and Speed on the output information.

We focus on a case study in which social networking site information into a GIS application that can be used by first responders, marketers etc. More precisely, we injected information from Twitter into World Wind. Twitter is one of the fastest growing online social networking services used to broadcast short messages (a.k.a. tweets), and World Wind is a GIS-based, Google Earth-like application developed by NASA.

The multiple sources of information were processed and analyzed simultaneously in near real-time and presented to users in an interactive and easily understandable format. The additional information was integrated into World Wind in such a way that it did not require the user to perform any additional activities other than the usual pan, zoom, and hovering needed to interact with a GIS application.

Next section provides additional background and related work, followed by a more detailed description of our case study and IQ dimensions in Section 3. Conclusions and Future work are presented in Section 4.

2. Related Work

Research into the use of short message data like Twitter to determine events (news) and their location is steadily increasing in popularity. Sankaranarayanan et al.

investigated the clustering of geo-tagged tweets with a spatial focus in their application called TwitterStand [4]. De Longueville et al adopted a direct form of geo-referencing Twitter data for tweet retrieval to aid in spatio-temporal information [5]. Phelan et al. proposed an approach of using real-time Twitter data as the basis for ranking RSS feeds to recommended news articles [6].

Furthermore, other researchers have examined some of the IQ issues of combining geo-spatial information with other categories of information. Hudson-Smith et al. [7] reviewed the development of various user designed functionalities that harnesses the power of web 2.0 in combining maps with social networking and other information available on the Internet and concluded that the value of information is increased by sharing it. A study conducted by Thakkar et al. [8] highlights the accuracy, representation and believability information quality challenges of the integration of geo-spatial information. Hariharan et al. [9] studied the accuracy and redundancy issues in maximizing within a given time limit the quality of GIS information from various sources.

Our research examines the combination of information from multiple sources in order to provide the user with a more enriched situational awareness without increasing the amount of activities/effort required by the user demand. The IQ aspects of the combination The IQ aspects in this paper are more complete and touch on multiple dimensions when compared to previous work.

2.1 *Geographic Information System and World Wind*

Geographic Information System (GIS) can be described as a system that can be used to manage and present location based data. It is essentially the digitization of static maps and other location based information. In contrast to the static maps, GIS allows users to interactively communicate, analyze and edit the data. Due to this dynamic and nature of the GIS, several applications that use geographically referenced data has been improved or developed e.g. GPS, remote sensing, and aerial photography.

World Wind is a free, open source virtual globe java application developed by NASA. It allows users to remotely access NASA, USGS, and publicly available GIS data such as satellite imagery, aerial photography, topographic maps, road maps and political boundaries which can each be used as different layers over the map [10].

We used World Wind as the GIS application in our case study due to its open source nature and the relative feasibility of creating a custom layer that could be used to present the additional information that we introduced into the GIS application.

2.2 *Social Networking Sites and Twitter*

Social networking sites are online services that allow individual to connect and exchange information with other people or groups that share common interests or attributes [11]. Several social networking sites that serve different purposes have sprung up in recent years. Facebook.com and Myspace.com are mostly used to stay in touch with friends, LinkedIn.com to connect with professional colleagues, Twitter.com for both keeping in touch with friends and broadcasting to the public. The popularity of social networking sites has increased tremendously over the years; together with blogs it is the fourth most popular internet activity and accounts for almost 10% of all internet time [12].

We chose Twitter as the social networking site for our case study mainly because of its microblogging feature. Since its creation in 2006, Twitter has become one of the largest social networking services available. In 2008, it grew at the rate of 752% for a total of 4.43 million unique visitors [13]. It allows users to send among a group of “follower” or broadcast short messages (a.k.a. tweets) which can be accessed via different media e.g. internet phone, custom application.

The volume of users broadcasting Tweets makes Twitter a source of vast amounts of various kinds of information. However, this magnitude of information created and broadcasted with little or no restriction other than the amount of characters contained in a tweet presents both opportunities and challenges to data mining efforts. It presents opportunities in the sense that it provides real time information about individuals, the public’s perspective on issues as well as current news events. Some of the problems that we perceive with this information are potential amount of “noise” created by wrong or unverifiable information.

Although Tom Anderson, a social media market researcher, described Twitter as a “Babylon of Spam” [14], it can also be argued that it is also a source of valuable (current) information. For example, Twitter played an important role in broadcasting information from within Iran during the Iranian election crises of 2009 [15]. Moreover, security forces used Twitter as a source of real time information during the Mumbai terrorist attack in 2008 [16], Fire department and weather monitoring organizations also provide updates to the public update via Twitter [17].

Our approach to integrating Twitter into a GIS environment enables us to access some of the information quality issues of Tweets such as believability and timeliness.

3. Case study: Geo-referencing Twitter Information on World Wind

We present a case study to demonstrate the IQ aspects of integrating Twitter into World Wind in order to create a more enriched situational awareness for the user. More specifically real time information from Twitter is logically and spatially displaced on a map so as to enable users to dynamically update their knowledge of a geographical region without having to manually - create queries to extract information from Twitter, sift through tweets, or acquire new skills for utilizing the Twitter-World Wind system.

We analyzed the results of our automatically generated Twitter queries for keywords (words with highest frequency), sentiment (mood expressed in the tweets e.g. panic, sadness and happiness), and trend (rate of keywords and sentiment occurrence over time relative to geographical locations).

Part of our motivation for using this case study is to examine the feasibility and possible shortfalls of using Twitter as a source of real-time information in the decision making process of first responders. In this paper we focus on the information quality (IQ) aspects of the study. We described the IQ issues using some of the data quality dimensions enumerated by Strong et al [18] namely timeliness, believability, accuracy, ease of use, and interpretability

3.1 Timeliness

In today's information technology driven society, timely and effective response to emergencies and natural disaster is critical to first responders, not only because of the lives and livelihood that have to be saved, but because it also helps to promote and maintain good public relation and sometimes to preserve employment. An example was FEMA's response to Hurricane Katrina in 2005.

We examined the lag time between the occurrence of an event and when indication of it appears in our keyword analysis. Suppose that had FEMA been able to better integrate information from the news media and bloggers[3] (most of the citizens learned about FEMA's inefficiencies though these media) with their other available information and displayed them spatially in a logical manner, it might have been possible to more effectively monitor and address the situations which contributed to the public backlash. Furthermore, the additional information might have saved more lives and livelihood.

We indentified two factors that contribute to the timely appearance of an event's keyword(s) on the map. These factors are the number of tweets regarding the event and the size of the geographical region affected/interested in the event. Arguably, an event is tweeted about within seconds of its occurrence therefore it is plausible to assume that the knowledge can be transferred and viewed immediately in the GIS environment.

However, in order for keywords that signal an event to show up on the map, not only must it be tweeted about but the number of tweets about it must be significant enough to give its signal/buzz words a relatively high frequency. Therefore, important events will naturally surface to the map and overcome day to day tweets and spam (See Figure 6). The manner in which the keywords are displayed also provides a sense of the geographical distribution of the event. Users can perceive whether an event is generated over a large geographical area or whether the event is concentrated into a single spot by flying down thereby narrowing down to specific areas of the map as shown in. Figure 5

Hence given an event like a hurricane, current trends indicate that within minutes the keywords about it would appear on the map. The LA fire department used Twitter to both inform and obtain from the public incidents of fire outbreaks [19] [20] .

The new information product creates a more effective situational awareness by providing timely information that could help the decision making process. We did not examine here the timeliness of other information such as place names and topography used in the GIS system , but that is likely to lag the timeliness of Twitter data because changes in place names or seasonal changes in landscape require some time to propagate, if ever, to the GIS databases.

[3] Twitter couldn't have been used during this period because it was not available until a year afterwards.

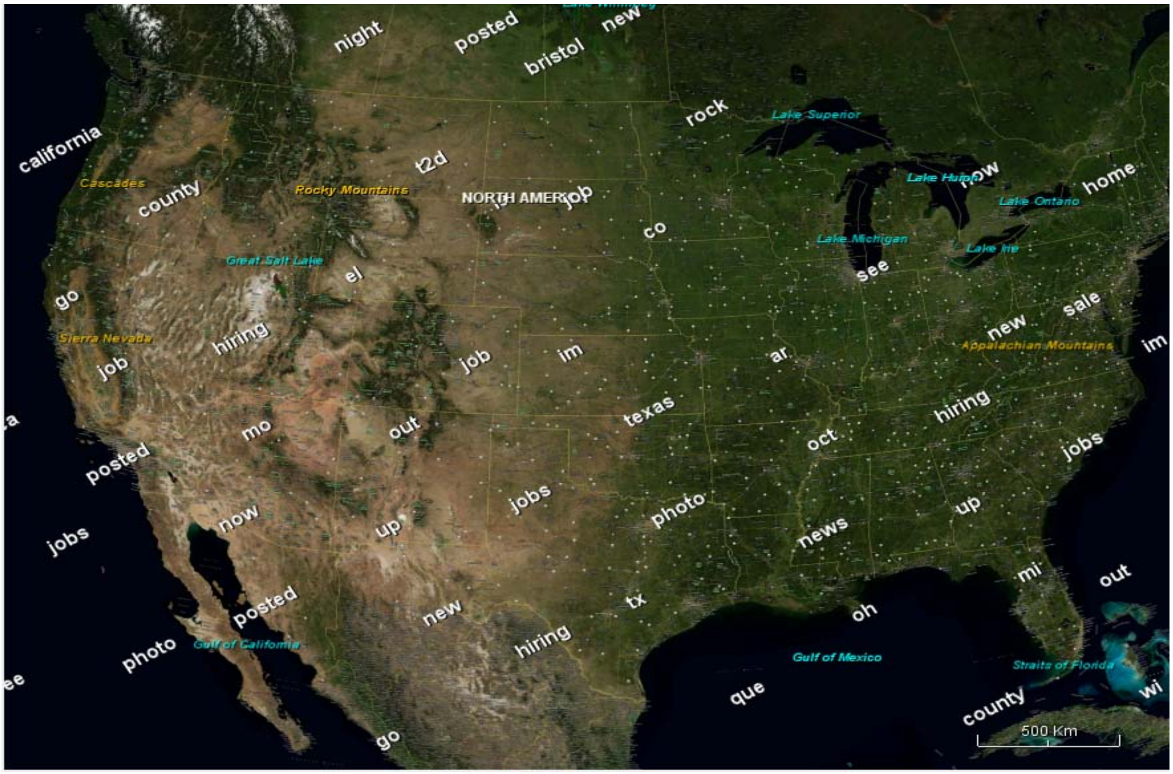


Figure 6: Continental USA Twitter data showing the most frequent words in tweets associated to the US

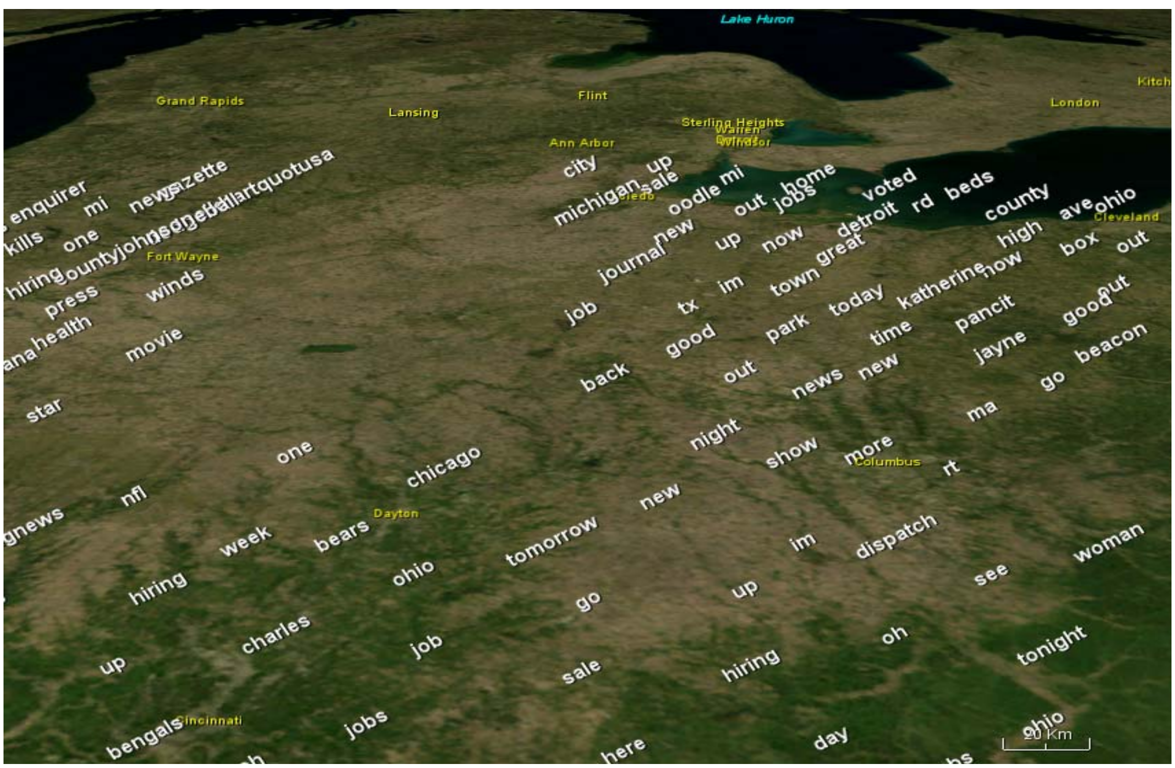


Figure 5: Regional Twitter data showing the most frequent words in tweets associated with Columbus and Dayton area

3.2 *Believability*

We examine under the believability dimension the perceived integrity of the information from Twitter. Because information can be broadcasted by anyone on Twitter due to the very little restrictions available, there is ample opportunity for misinformation to be perpetuated. For this reason and because of the potentially critical decisions that would be based on our information product, believability is a very critical issue. To put it briefly, if users do not believe the information presented by the system, they would not rely on it for decision making, hence their decision making process is likely to remain the same or more complicated due to possible extra complexity introduced by Twitter.

NASA and USGA GIS data available on World Wind have a high level of integrity, hence they are believable. The integration of Twitter into the GIS environment effectively creates an information product with both "high" and "low" believability. This integration can be used as an opportunity to increase the believability of the information derived from Twitter in certain cases. Essentially, by putting Twitter information in the context of location context its integrity can be more easily discernable and checked For example, if a collection of tweets indicates the occurrence of flooding in the Mojave Desert, one may use this reason to easily conclude that such an event is not likely. Furthermore, if multiple contradictory events occur in the same location or neighboring locations the user is better positioned to determine the believability of the information.

Generally, the more "reliable" information available to disambiguate the "less reliable" information, the easier it is to determine the believability of the "less reliable" information.

3.3 *Accuracy*

Under the accuracy dimension we analyzed how correctly the analyses of the tweets were displayed geo-spatially. The three factors we identified as affecting the preciseness of the geographical placement of keywords relative to the associated location are the focus altitude, proximity of places and the number of words displayed.

The focus altitude is the distance above the sea level between the user's view and the earth's surface. Because altitude is a determining factor in the number and size of tiles used by World Wind to display surfaces, geo-referencing keywords is consequently affected. . In effect, the same screen space is used to display information regardless of whether the view presents the whole United States or just a single city. In the US view, the placement of Twitter keywords is less precise than in the view of the city because we take into account not only the ideal position of a keyword, but also its potential readability. As such, keywords may need to be moved around to provide enough inter-word spacing

At very low altitudes, the tiles are small enough such that individual town/city on the PlaceName layers can be displayed on individual tiles. When places have proximity to one another, their displayed keywords sometimes overlap creating the possibility for wrong associations or confusions. The second and third factors are somewhat related because the overlap in keywords for different places is dependent on the number of keywords displayed for each place.

The accuracy of our text analysis i.e. analysis of tweets for keywords, sentiment and are beyond the scope of this paper.

3.4 *Ease of Use*

Since many first responders' (e.g. military, red cross) command and control centers already use location based information in their operations, integrating new information in the context of location relatively helps in its assimilation. Furthermore, serving up Twitter information on an existing and familiar platform to the users concerned is preferable to having them master the intricacies of new applications. This can be very important when the time between learning to use new applications and responding to is very small.

In addition to presenting the new information on an existing platform (GIS), the ease of use of the information was also examined from the perspective of the amount of additional user activities/efforts required to operate the new system. It is therefore desirable that the complexity of the system from this point is not increased significantly. Although the more technical details are not published here, the user is not required to do more than the usual pan, zoom and hovering needed to maneuver on World Wind. Twitter query is dynamically generated and the tweets are automatically analyzed and geo-referenced on the map.

The exploration of Twitter keywords in the GIS context is in fact completely free for the user, and no new skills are required. Navigating through the Earth results in automatic filtering and aggregation or drill-down of the Twitter data.

3.5 *Interpretability*

This IQ dimension is improved for Twitter data because the user can at-a-glance see the most prominent keywords, correlate tweets with GIS, and discover new relationships between sets of tweets.

The display of the keywords shows information extracted from hundreds of tweets over an easy to interpret geographical milieu. The task of reading each tweet individually and of understanding the overall structure requires significantly more resources from a user. Furthermore, even alternate forms of displaying the extracted keywords, such as tables, would still be harder to interpret and navigate (for example drill-down or increase the level of aggregation) than a map.

Another interpretability boost stems from displaying Twitter data in the context of a map. It is not unusual to have incomplete information in a tweet, sometimes by omission and sometimes because that information is self evident to the sender. For example, there may be tweets that talk about an accident on the interstate south of XYZ. Without a map, it may be impossible to determine which interstate has the accident, but GIS data can simply disambiguate the highway.

Finally, placing keywords next to each other uncovers relationships between tweets that would be difficult to observe otherwise. Users can see events occurring in adjacent cities, states, and even countries. The spatial placement provides links between keywords that are not intrinsically written in tweets. User can, for example, compare and contrast events taking place in Central Arkansas with events in the Fayetteville area.

4. Conclusions

We presented a case study that demonstrates the use of tweets in a GIS environment to increase situational awareness. We discussed the information quality aspects of the combination of these two categories of information under the IQ dimensions of the timeliness of tweets enhances GIS data, while the increased believability, accuracy, ease of use and interpretability of geographical data is exported to Twitter We highlighted the advantage of using location information to assess the quality of information derived from tweets. Generally, we explored a new use for textual data in an environment for which it was not originally intended; this maximization of its use concurs with one of the philosophies of information quality,

Further research is currently being conducted to assess user behavior and interaction with the new information product. Some of the aspects under research include the effectiveness of the tweets analysis in capturing events of different magnitudes, introduction of more factors in the dynamic generation of tweets query.

Acknowledgments

The research for this paper was supported in part by the US Air Force Research Laboratory Sensor and Human Factors Directorates at Wright Patterson Airbase (SRA Contract #S280000131). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of any of the above organizations or any person connected with them.

References

[1] H. F. Korth. and A. Silberschatz. *Database research faces the information explosion.* 2, s.l. : ACM New York, NY, USA, February 1997, Communications of the ACM, Vol. 40, pp. 139-142.

[2] J. R. Talburt.*A New View of Information Quality.* Seoul, South Korea : s.n., 2009. Database Grand Conference. pp. 241-251.

[3] Drori, Offer.*Using an information reduction model in hypertext virtual node as a direction for solving the data explosion problem.* 3, s.l. : ACM New York, NY, USA , December 1995, ACM SIGWEB Newsletter , Vol. 4, pp. 10-14.

[4] J. Sankaranarayanan, et al.*TwitterStand: News in Tweets.* Seattle : ACM, 2009. The 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. pp. 42-51.

[5] B. De Longueville, R. .S. Smith. and G. Luraschi.*"OMG, from here, I can see the flames!": a use case of mining Location Based Social Networks to acquire spatio-temporal data on forest fires.* Seattle : ACM, 2009. The 2009 International Workshop on Location Based Social Networks. pp. 73-80.

[6] O. Phelan, K. McCarthy and B. Smyth.*Using Twitter to Recommend Real-Time Topical News.* New York : ACM, 2009. The 3rd ACM Conference on Recommender Systems. pp. 385 - 388.

[7] A. Hudson-Smith.*Mapping for the Masses.* s.l. : Sage Publications, 2009, Social Science Computer Review OnlineFirst, pp. 1-15.

[8] S. Thakkar, C. A. Knoblock and J. L. Ambite. *Quality-Driven Geospatial Data Integration.* Seattle : ACM , 2007. The 15th International Symposium on Advances in Geographic Information Systems. pp. 1-8.

[9] Hariharan, Ramaswamy, et al.*Quality-Driven Approximate Methods for Integrating GIS Data.*

Bremen : ACM, 2005. The 13th Annual ACM International Workshop on Geographic Information Systems. pp. 97-104.

[10] NASA. World Wind. *World Wind.* [Online] [Cited: February 9, 2010.] http://ti.arc.nasa.gov/projects/worldwind/index.php.

[11] D. M. Boyd and N.B. Ellison. *Social network sites: Definition, history, and scholarship.* 1, 2007, Journal of Computer-Mediated Communication, Vol. 13.

[12] Neilsen Corp. *Global Faces and Networked Places-A Nielsen report on Social Networking's New Global Footprint.* s.l. : The Nielsen Company, 2009.

[13] A. Ostrow. Twitter's Massive 2008: 752 Percent Growth. *Mashable The Social Media Guide.* [Online] [Cited: February 9, 2010.] http://mashable.com/2009/01/09/twitter-growth-2008/.

[14] T. H. Anderson. Reward for Being the Top Market Researcher on Twitter. *Tom H. C. Anderson - Next Gen Market Research.* [Online] Next Gen Market Researcher, February 3, 2010 . [Cited: February 4, 2010.] http://www.tomhcanderson.com/2010/02/03/reward-of-being-the-top-market-researcher-on-twitter/.

[15] E. Morozov. Iran Elections: A Twitter Revolution? *The Washington Post.* [Online] June 17, 2009. [Cited: February 9, 2010.] http://www.washingtonpost.com/wp-dyn/content/discussion/2009/06/17/DI2009061702232.html.

[16] C. Beaumont. Mumbai attacks: Twitter and Flickr used to break news. *Telegraph.co.uk.* [Online] Novomber 27, 2008. [Cited: February 9, 2010.] http://www.telegraph.co.uk/news/worldnews/asia/india/3530640/Mumbai-attacks-Twitter-and-Flickr-used-to-break-news-Bombay-India.html.

[17] H. Havenstein, (Computerworld). LA Fire Department all 'aTwitter' over Web 2.0. *PC World.* [Online] August 3, 2007. [Cited: February 9, 2010.] http://www.pcworld.com/article/135518/la_fire_department_all_atwitter_over_web_20.html.

[18] D. M. Strong,Y. W. Lee, Yang W. and R.Y.Wang. *Data Quality in Context.* 5, s.l. : ACM Publication, May 1997, COMMUNICATIONS OF THE ACM, Vol. 40, p. 8.

[19] H. Havenstein. LA Fire Department all 'aTwitter' over Web 2.0. *PCWorld.* [Online] August 3, 2007. [Cited: February 15, 2010.] http://www.pcworld.com/article/135518/la_fire_department_all_atwitter_over_web_20.html.

[20] D. Tabor. LAFD's One-Man Geek Squad Brings Web 2.0 to Firefighting. *WIRED MAGAZINE.* [Online] October 20, 2008. [Cited: February 15, 2010.] http://www.wired.com/entertainment/theweb/magazine/16-11/st_firefight.

[21] B. Cahoon,K. S. McKinley and Z.Lu. *Evaluating the performance of distributed architectures for information retrieval using a variety of workloads.* 1, s.l. : ACM New York, NY, USA , January 2000, Transactions on Information Systems (TOIS), Vol. 18, pp. 1-43.

[22] www.wantoknow.com. FEMA Failures: Hurricane Katrina (FEMA Failures Reported in Major Media). *www.wantoknow.com.* [Online] [Cited: February 9, 2010.] http://www.wanttoknow.info/femafailureskatrina.

APPENDIX 3: TRACK 3 - VISUAL RENDERING AND DISPLAY OF TEXT & IQ METRICS – SMART ENVIRONMENT

Appendix 3-A: "Bayesian Data Fusion for Smart Environments with Heterogenous Sensors"

"Bayesian Data Fusion for Smart Environments with Heterogenous Sensors" Soukaina Messsoudi, Kamilia Messaoudi, Serhan Dagtas – University of Arkansas at Little Rock. Presented at the 8th Annual Consortium for Computing Sciences in Colleges; Searcy, Arkansas; March 26-27, 2010.

Abstract *- Smart environments refer to buildings or locations equipped with a multitude of sensors and processing mechanisms for improved security, efficiency or functionality. Often, these sensors serve distinct purposes and their data may be processed separately by entirely separate systems. We argue that integrated processing of data available from multiple types of sensors can benefit a variety of decision making processes. For example, smart building sensors such as occupancy or temperature sensors used for lighting or heating efficiency can benefit the security system, or vice versa. Recent industry standards in sensor networks such as ZigBee make it possible to collect and aggregate data from multiple, heterogeneous sensors efficiently. However, integrated information processing with a diverse set of sensor data is still a challenge. We provide an information processing scheme that offers data fusion for multiple sensors such as temperature sensors or motion detectors and visual sensors such as security cameras. The broader goal of multi-sensor data fusion in this context is to enhance security systems, improve energy efficiency by supporting the decision making process based on relevant and accurate information gathered from different sensors. In particular, we investigate a major data fusion technique, Bayesian network, and present a simulation tool for a "smart environment". In addition, we discuss the potential impact of data fusion on the processes of decision or detection, estimation, association, and uncertainty management.*

BAYESIAN DATA FUSION FOR SMART ENVIRONMENTS with HETEROGENOUS SENSORS

Soukaina Messaoudi, Kamilia Messaoudi and Serhan Dagtas
Department of Information Science
University of Arkansas at Little Rock
2801 S. University Avenue
Little Rock AR 72204
sxmessaoudi@ualr.edu, kxmessaoudi@ualr.edu, sxdagtas@ualr.edu
(501)683-7267

ABSTRACT: *Smart environments refer to buildings or locations equipped with a multitude of sensors and processing mechanisms for improved security, efficiency or functionality. Often, these sensors serve distinct purposes and their data may be processed separately by entirely separate systems. We argue that integrated processing of data available from multiple types of sensors can benefit a variety of decision making processes. For example, smart building sensors such as occupancy or temperature sensors used for lighting or heating efficiency can benefit the security system, or vice versa. Recent industry standards in sensor networks such as ZigBee make it possible to collect and aggregate data from multiple, heterogeneous sensors efficiently. However, integrated information processing with a diverse set of sensor data is still a challenge. We provide an information processing scheme that offers data fusion for multiple sensors such as temperature sensors or motion detectors and visual sensors such as security cameras. The broader goal of multi-sensor data fusion in this context is to enhance security systems, improve energy efficiency by supporting the decision making process based on relevant and accurate information gathered from different sensors. In particular, we investigate a major data fusion technique, Bayesian network, and present a simulation tool for a "smart environment". In addition, we discuss the potential impact of data fusion on the processes of decision or detection, estimation, association, and uncertainty management.*

Key Words: Data and information fusion, Bayesian, Dempster-Shafer, Fuzzy logic, Neural Networks, Visual sensors, Non-visual sensors, Sensor networks, Motion segmentation, OpenCV.[1]

INTRODUCTION

One of the outcomes of data fusion is the improved information quality that assists various decision making processes in a "smart environment". Our focus here is the integration of sensors information into the real-time decision making process in a surveillance context. We use data fusion in a fashion where different types of information are collected from a heterogeneous set of visual and non-visual sensors. The process of integrating data from different sources requires designing an appropriate data fusion model that would take the sensor data, integrate them following a certain model, and transform it to a set of useful and relevant decisions. The anticipation is for the resulting decisions to be more accurate and efficient than those resulting from a single source. In a broader sense, we expect data fusion to lead to a virtual collaboration between the different collected information.

Towards this goal, we first investigate the usefulness of data fusion in a smart environment equipped with visual and non-visual sensors and design a convenient data fusion model. Then, we provide an overview of data fusion methods, present our data fusion algorithm and discuss our data fusion engine. This is followed by a description of our smart environment simulation tool which is used to test some of the hypotheses, visualize the environment with the sensors and their spatial relationships

Distribution A; Approved for Public Release; Distributed Unlimited; 88ABW/PA cleared 24 September 2012 as 88ABW-2012-5092.

and to allow us to build some of the case scenarios which is discussed last. In the last section, we summarize our findings and conclusions with a set of ideas for ongoing work.

TECHNIQUES FOR DATA FUSION

Data fusion is "the theory, techniques and tools which are used for combining sensor data, or data derived from sensory data, into a common representational format." Fusing data from different sources can improve the quality and the utility of information and help improve efficiency, security and functionality. The critical problem in multi-sensor data fusion is to determine the best procedure for combining information from different sensors in the system.

Most of the reported work in data fusion uses a statistical approach in order to describe different relationships between sensors taking into account the underlying uncertainties [4]. Edward Waltz and James Llinas summarize the methods to implement data fusion as follows: decision or detection, estimation, association, and uncertainty management theories. In decision or detection theory "measurements are compared with alternative hypotheses to decide which ones best describe the measurement." Basically, the decision theory assumes "the probability descriptions of the measurement values and prior knowledge to compute a probability value for each hypothesis." [2].

Fuzzy logic, neural networks, Bayesian, and Dempster-Shafer theories are the most commonly used methods in multi-sensor data fusion. However, our approach will focus on Bayesian model for integrated information processing using data from multiple, heterogeneous sensors. The main reasons for this election were the appropriateness of the input and output types in Bayesian model and its wide-spread use for similar problems in the literature. We plan to expand our work into the alternative fusion techniques as part of our ongoing research.

The basic principle of *Bayesian theory* is that all the unknowns are treated as random variables and that the knowledge of these quantities can be represented by a probability distribution. In addition, Bayesian methodology claims that the probability of a certain event represents the degree of belief that such an event will happen. The degree of belief is associated with a probability measure that can be updated by additional observed data. All the new observations are added to update the prior probability and therefore obtain a posterior probability distribution [3].

BAYESIAN DATA FUSION

The Bayesian model integrates data, independently, from **r** correlated sensors' inputs in the following pattern:

$$\mathrm{p}(\mathrm{D}/\, X_1^1 X_1^2 \ldots X_1^r) = \frac{\prod_{j=1}^{r} p(D/X_1^j) * p(D/X_0^1 X_0^2 \ldots X_0^r)}{\prod_{j=1}^{r} p(D/X_0^j)} * \mathrm{K}$$

where K is the Bayesian normalization and is equivalent to $\frac{\prod_{j=1}^{r} p(x_1^j/X_0^j)}{p(x_1^1 x_1^2 \ldots x_1^r / X_0^1 X_0^2 \ldots X_0^r)}$ and p(D/ $X_1^1 X_1^2 \ldots X_1^r$) is the probability of event D given $X_1^1, X_1^2, \ldots, X_1^r$.

x_1^j: Current measurement/observation from correlated sensors j where j = 1, 2, …,r. X_0^j: Prior information or old data set from correlated sensors j where j = 1, 2, …,r. X_1^j: Posterior information or new data set from correlated sensors j where j = 1, 2, …,r. D : Event in question (one of the decisions labeled on figure1).

The fusion engine in this project is the model we use to integrate information from both visual and non-visual sensors. The engine we design receives inputs from both visual and non-visual sensors and provides a set of relevant decisions (outputs).

As the diagram in Figure 1 shows, s_1, s_2, $s_3 \ldots s_n$ are inputs from different non-visual sensors. These inputs first go through a correlation model (raw data processing on figure1) that determinates the correlations among the sensors' inputs and transmits independent *m* outputs that are fed to the fusion engine as inputs. These outputs (fusion engine inputs) are labeled as x_1, x_2, $\square_3 \ldots \square_{\square}$.

The fusion engine inputs $\square_1$, $\square_2$, $\square_3 \ldots \square_{\square}$ can be matched to notations such as $\square_I^1$, $\square_I^2$, $\square_I^3 \ldots$, $\square_I^{\square}$, which represent the posterior information, described in the algorithm section, from correlated sensors. However, this matching does not restrict matching $\square_1$ to $\square_I^1$, $\square_2$ to $\square_I^2 \ldots$etc as the data fusion model we use consider integrating posterior information from both non-visual and visual sensors. As it is explained below, data from visual sensors is pre-processed before it can be fed to the fusion engine. This pre-processing results in a convenient format of information to be passed to the fusion engine.

For visual sensors, we use optical and infrared cameras to record raw videos. The acquired videos are then processed to extract meta-data information to be used in the fusion algorithm described above. The processing of images from such visual sensors requires a preliminary processing where some intermediary image features such as moving objects and their boundaries are extracted for further processing [5]. The final extracted visual information forms metadata that can be fed to the designed fusion engine that integrates it with other sensor data from other heterogeneous sensors.

The extraction of visual information can be a real challenge because of "the lack of proper low-level algorithms for robust feature extraction" [7]. Here, we use a motion detection algorithm to extract relevant visual information about the moving objects in the recorded video. The algorithm chosen for this purpose is the implementation in OpenCV, which is an open-source computer vision library, originally developed by Intel. We have performed a few modifications at the input level that resulted in movement detection. The metadata in this context includes the kind of information such as the number of moving objects, the nature of movement, the type of the moving objects (human or animal), the actions performed by the moving objects, the area they occupy, and the time they stay in the room of question.

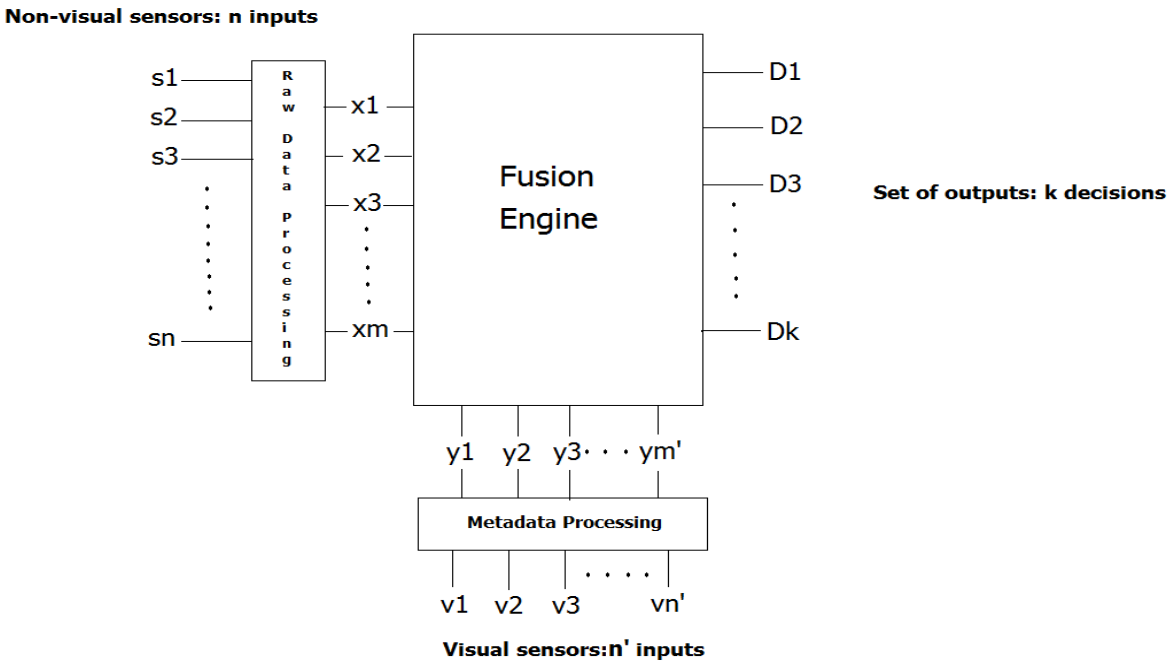


Figure1: Fusion Engine Design

In the fusion engine design on Figure1, $\square_{1}$, $\square_{2}$, $\square_{3}$…, $\square_{\square'}$ represent the information collected (metadata) from the every visual sensor (n' visual sensors). These inputs are processed (metadata processing in Figure 1) to create appropriate input format. The resulting outputs of the metadata processing are also in the form of correlated information. In other words, some visual sensors can be correlated in the sense that only one output can be retrieved from them. This correlation of visual sensors results in independent inputs labeled as $\square_{1}$, $\square_{2}$, $\square_{3}$…, $\square_{\square'}$ in Figure 1.

After tracking moving objects on a given video, more work is done on detecting the different features of these moving objects. Features such as the number of moving objects, the nature of the moving objects (human, animal…), and the nature of movements (fast, slow…) the objects perform are examples of information we want to feed to the fusion engine. After extracting such important information (metadata), we perform another processing on the metadata to come up with an input format compatible with the data fusion model we are using (Bayesian model).

In data fusion context, the outputs of such a model are in the form of decisions that should be performed to better serve the environment where the different types of sensors are used. As Figure 1 shows, the set of decisions D1, D2,…, Dk are the independent fusion engine outputs (or decisions). These decisions can help in saving energy, restricting security, launching rescue operations and many more. Depending on what type of sensors we use, a set of relevant and efficient decisions can be formed.

SIMULATION TOOL & EXPERIMENTS

In our study of multi-sensor data fusion, we implement a simulation tool that helps us construct a virtual smart environment. The smart environment has basically different types of sensors such as: motion detector, smoke detector, daylight sensor, and other types of sensors. In addition to sensors, there are objects that can be moving around to generate case scenarios where motion is a factor to be considered. Emergency cases such as fire or flood can be studied using the implemented simulation tool. This tool is implemented using JAVA and it facilitates the study of multiple scenarios because the user can chose any type of sensors implemented in the tool as well as manage the environment's state such as increasing the temperature (fire case) or adding moving objects or water (flood scenario). Visual sensors are placed on the simulation grid at specific grid locations. A specific set of attributes must be defined for each sensor. These may include range, angle, sensitivity, and direction. Every sensor has a detection area and detection occurs when the coverage area and attributes of a given object overlap with the detection range and sensitivities of a given sensor. The simulation tool is our main data generator where sensors' flags and data are fed to the fusion engine where decision making process takes place.

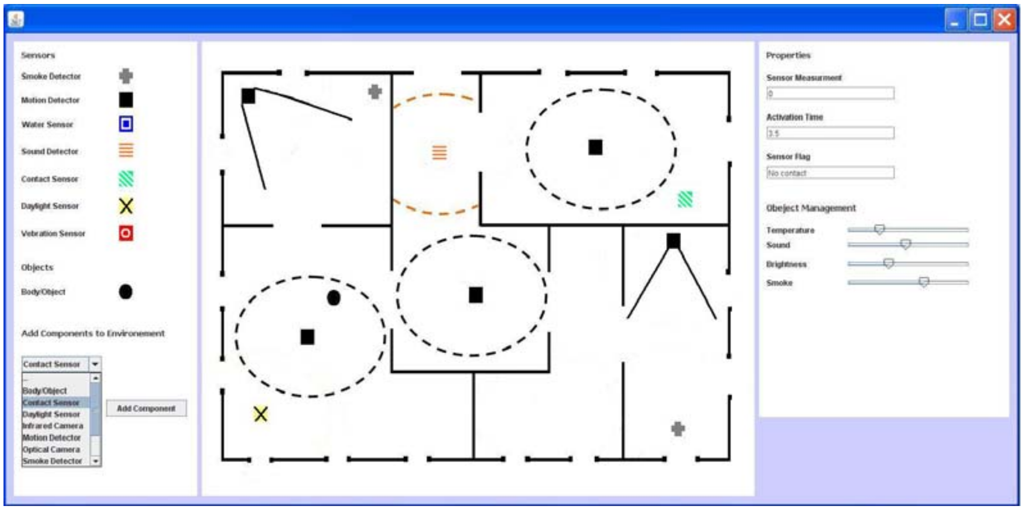


Figure 2: Simulation tool interface

In order to develop a reasonable method for finding a likelihood function at a given moment, we have carefully studied the behavior of the moving objects. We have conducted ten experiments where we tracked one object in every video and recorded the corresponding data. Because of space limitations, we present only the conclusions we have deived from the analysis of data. Through analyzing the graphs from the experiments, we take into consideration the factor of persistence, which merely means for how long the object(s) is moving. In order to do this, we choose a time instance from the plot and study the behavior of the moving object in previous time instants.

The probability value that will be used by the fusion engine at time $\square_{\square}$ is computed as follows:

$\square_{\square} = \sum_{\square=0}^{\square-1} \square(\square_{\square}).[\frac{1-\square}{10}]$ where $\square(\square_{\square})$ is $\square_{\square}$'s equivalent area percentage

In order to find the reasonable number of previous time instants that should be included in the computation of a given likelihood function at a given instant, we further analyze the data collected from the ten experiments. We apply the formula above at $\square_9$ for every experiment and find the equivalent likelihood function ($\square_9$) taking into consideration m=10, 8, 6, and 4 previous time instants.

The table below summarizes the analysis:

	Exp1	Exp2	Exp3	Exp4	Exp5	Exp6	Exp7	Exp8	Exp9	Exp10
m=10	117.70%	108.40%	75.83%	210.48%	62.10%	27.29%	84.31%	195.53%	72.45%	80.17%
m=8	77.50%	77.84%	55.21%	151.95%	40.81%	20.20%	58.00%	127.93%	48.65%	51.80%
m=6	56.93%	48.38%	35.18%	95.94%	22.13%	13.10%	34.40%	76.18%	28.03%	28.60%
m=4	28.60%	24.66%	18.76%	52.01%	9.48%	6.52%	16.17%	36.50%	13.81%	12.82%

Table 1: Computed likelihood function at □□ using the weighted method.

From the table above, we conclude that looking back at eight or six time instants usually result in a reasonable value that gives us an idea about how intense the motion is in a given room and can safely be fed to the fusion engine. Also, the computation of a likelihood function for m=8 or 6 is easy and quicker than m=10 or more; it also doesn't take into consideration the percentage value at t_0 where usually no motion is recorded.

CONCLUSIONS

We have demonstrated ways to use Bayesian data fusion technique in a smart environment with a heterogeneous, inter-dependent set of sensors. This was done by generating statistically independent inputs for the Bayesian fusion model and demonstrate the effect through a simulation tool. The *Dempster-Shafer theory* is considered to be a generalization of the Bayesian theory of subjective probability. Dempster-Shafer allows us to "base degrees of belief for one question on probabilities for a related question" [6]. One of the most important advantages of the Dempster-Shafer theory is that it does not associate probabilities to questions of interest as Bayesian methods do. Instead, the belief for one question is based on probabilities for a related question; therefore, the Dempster-Shafer theory can effectively model uncertainty. As a next step, we plan to build a Dempster-Shafer model and draw comparisons with the Bayesian model. Additionally, further experimentation is underway using a testbed created by ZigBee-based sensors that implement the smart environment and by optical and infra-red cameras.

REFERENCES:

[1] Challa, S., Koks, D., An Introduction to Bayesian and Dempster-Shafer Data Fusion, http://robotics.caltech.edu/~jerma/research_papers/BayesChapmanKolmogorov.pdf, retrieved October 28, 2009.

[2] Waltz, Edward, James, L. ,*Multisensor Data Fusion*. Norwood: Artech House, Inc, 1990.

[3] Fisher C., Eitel L., Richard W., and Shobha C., *Introduction to Information Quality*. Cambridge : MITIQ, 2006.

[4] H.B, Mitchell, *Multi-Sensor Data Fusion*. Berlin: Springer, 2007.

[5] Snidaro, L., Claudio, P., Christian, M.,Gian, F., Visual Sensor Technology for Advanced Surveillance Systems: Historical View, Technological Aspects and Research Activities in Italy, http://www.mdpi.com/1424-8220/9/4/2252/pdf, retrieved November 12, 2009.

[6] Shafer, G., Dempster-Shafer Theory, http://www.glennshafer.com/assets/downloads/articles/article48.pdf, retrieved October 28, 2009.

Appendix 3-B: "Efficient Information Process in Smart Environments with Heterogenous Wireless Sensor Networks"

"Efficient Information Process in Smart Environments with Heterogenous Wireless Sensor Networks" Soukaina Messaoudi, Kamilia Messaoudi, Serhan Dagtas – University of Arkansas at Little Rock. Presented at the 2010 ALAR Conference on Applied Research in Information Technology; Conway, Arkansas; April 9, 2010.

Abstract *- Smart environments are buildings or locations equipped with a multitude of sensors and processing mechanisms for improved security, efficiency or functionality. Integrated processing of data available from multiple types of sensors can benefit a variety of decision making processes. Recent industry standards in sensor networks, such as ZigBee, make it possible to collect and aggregate data from multiple, heterogeneous sensors efficiently. The optimal design and placement of sensors in a two-dimensional or three-dimensional space, however is still an important challenge. In addition, integrated information processing with a diverse set of sensor data is another important research field. We present these two important cases with references to specific applications and discuss some possible solutions. The specific goal of integrated, efficient information processing in smart environments in such contexts is to enhance security systems, improve energy efficiency by supporting the decision making process based on relevant and accurate information gathered from different sensors. In particular, we investigate the use of Dempster-Shafer based data fusion model and present techniques for processing of visual sensor data to facilitate the use of Dempster-Shafer model.*

Efficient Information Processing in Smart Environments with Heterogeneous Wireless Sensor Networks

Kamilia Messaoudi
Department of Information Science
Univ. of Arkansas at Little Rock
2801 S. University Avenue
Little Rock AR 72204
kxmessaoudi@ualr.edu

Soukaina Messaoudi
Department of Information Science
Univ. of Arkansas at Little Rock
2801 S. University Avenue
Little Rock AR 72204
sxmessaoudi@ualr.edu

Serhan Dagtas
Department of Information Science
Univ. of Arkansas at Little Rock
2801 S. University Avenue
Little Rock AR 72204
sxdagtas@ualr.edu

Abstract

Smart environments are buildings or locations equipped with a multitude of sensors and processing mechanisms for improved security, efficiency or functionality. Integrated processing of data available from multiple types of sensors can benefit a variety of decision making processes. Recent industry standards in sensor networks, such as ZigBee, make it possible to collect and aggregate data from multiple, heterogeneous sensors efficiently. The optimal design and placement of sensors in a two-dimensional or three-dimensional space, however is still an important challenge. In addition, integrated information processing with a diverse set of sensor data is another important research field. We present these two important cases with references to specific applications and discuss some possible solutions. The specific goal of integrated, efficient information processing in smart environments in such contexts is to enhance security systems, improve energy efficiency by supporting the decision making process based on relevant and accurate information gathered from different sensors. In particular, we investigate the use of Dempster-Shafer based data fusion model and present techniques for processing of visual sensor data to facilitate the use of Dempster-Shafer model.

Keywords: Smart environment, wireless sensor network, Dempster-Shafer, Bayesian, data fusion, controllers, system optimization, information processing.

1. Introduction

Smart environment refers to any open or closed field equipped with hydrogenous sensors, actuators, displays, and computational elements connected through a continuous wireless network. Usually, smart environments provide the ability to automate the environment and replace the physical labor with automated agents. There are three different types of smart environments in Poslad's point of view: virtual computing environments, physical environments and human environments, or a combination of all previously listed types [7].

Smart environments have many features such as remote control of devices, device communication, information acquisition and dissemination from sensor networks, enhanced services by intelligent devices, and predictive and decision-making capabilities. Technologies used in smart environments involve wireless communication, adaptive control, parallel processing, image processing, image recognition, signal prediction and classification, sensor design, motion detection, and many others.

Our focus in this paper is the integrated, efficient processing of sensor information in a smart environment from the viewpoints of information quality, data fusion and network efficiency. We present the prominent issues in each of these fields and several aspects of our integrated, data-fusion-based approach are discussed in the next section.

2. Information Processing in Smart Environments

As widely accepted, smart environments rely on sensory data from the real world as humans do. The sensory data comes from multiple sources or sensors of different modalities in distributed locations, i.e. Wireless Sensor Networks [10]. As a result, many challenges that concern "detecting the relevant quantities, monitoring and collecting the data, assessing and evaluating the information, formulating meaningful user displays, and performing decision-making and alarm functions" develop and need to be dealt with correctly using WSNs. Information needed in a smart environment can be best provided by a distributed wireless sensor network that are responsible for sensing and further processing of information. This is the reason wireless sensor networks have gained much importance recently and more research is being conducted on this area, especially in the processing of heterogeneous data effectively and efficiently.

A challenging issue in the smart environments is the restricted quality of sensory data that is due to sensor failures or limited precisions [8]. Data stream processing introduces noise and decreases the data quality in streaming environments. As a result, a lot of business decisions can be wrong since they are based on dirty or simply wrong data. In order to avoid this issue, quality characteristics need to be captured and provided to the business task.

Distribution A; Approved for Public Release; Distributed Unlimited; 88ABW/PA cleared 24 September 2012 as 88ABW-2012-5092.

Another challenge in information processing concerns securing information in a distributed wireless sensor network. In [11], the authors address the issue of securing in-network processing for wireless sensor networks and suggest mechanisms for securing both upstream data aggregation and downstream data dissemination.

Taking all these issues into consideration, much research has been done to improve the security and privacy in wireless sensor networks and to enhance the mechanisms used in processing the information received from the multiple sensors of the network. Data fusion is a concept that goes hand in hand with wireless sensor networks as it tends to improve the quality of information retrieved from such a network. Below, we discuss several data fusion concepts and the prominent methods used to integrate data from a multi-sensor network.

Data Fusion in Wireless Sensor Networks (WSN)

The outcomes of data fusion in the smart environment context are related to the improvement of information quality that assists various decision making processes. In our previous research about data fusion and the integration of sensors information into the real-time decision making, we presented a case where data is collected from heterogeneous set of visual and non-visual sensors [18]. The process of integrating data from different sources requires designing an appropriate data fusion model that would take the sensor data, integrate them following a certain model, and transform it to a set of useful and relevant decisions. The anticipation is for the resulting decisions to be more accurate and efficient than those resulting from a single source. In general terms, *data fusion* is "the theory, techniques and tools which are used for combining sensor data, or data derived from sensory data, into a common representational format." However, critical problem in multi-sensor data fusion is to determine the best procedure for combining information from different sensors in the system [13].

Most of the reported work in data fusion uses a statistical approach in order to describe different relationships between sensors taking into account the underlying uncertainties [13]. Edward Waltz and James Llinas summarize the methods to implement data fusion as follows: decision or detection, estimation, association, and uncertainty management theories. In decision or detection theory, "measurements are compared with alternative hypotheses to decide which hypothesis best describes the measurement." Basically, the decision theory assumes probabilistic descriptions of measured values and prior knowledge in order to compute a probabilistic value for every hypothesis [8]. Association occurs when the fusion system uses multiple measurements from different sources and they must be associated with each other prior to, or at least in conjunction with, a classification or estimation. The correlation process should be performed to quantify a measure of the correlation among all measurements in order to partition measurements into sets. The measurements in each set are associated with a common source.

Uncertainty management stems from classical methods that represent uncertainty in measurements using the Bayesian probability model to express the degree of belief in each hypothesis as a probability. The hypothesis must be mutually exclusive and this requires that all hypotheses must form a complete set of possibilities and the probabilities must sum to one. Because the Bayesian model cannot present uncertainty along with the fact that probabilities must be assigned to each hypothesis, Dempster-Shafer introduced the concept of probability intervals to provide means to express uncertainty. Other heuristic models and fuzzy calculus have also been applied to uncertainty representation for fusion applications [8].

Our approach focuses on integrating information from different sensors, which, we argue, is best accomplished by data fusion. We have devised Bayesian and Dempster-Shafer based data fusion models and compared the results from both methods. A design of the fusion engine that can be either Bayesian or Dempster-Shafer is provided below (Figure 2). This figure explains that data is retrieved from different kinds of sensors, visual and non-visual, which depend on the application being used at. Data from every sensor should first go through a preliminary data processing where collaborations among sensors are determined and likelihood functions (Bayesian model) or mass functions (Dempster-Shafer model) are formed.

In Figure 2, $\square_1, \square_2, \square_3 \ldots \square_{\square}$ represent inputs or readings from different non-visual sensors. These inputs are further processed to formulate $\square_1, \square_2, \square_3 \ldots \square_{\square}$, which represent the appropriate likelihood functions (probabilities for Bayesian model) or mass functions, which are sensors' beliefs that should be fed to the Dempster-Shafer model. In a similar manner, $\square_1, \square_2, \square_3 \ldots, \square_{\square'}$ represent metadata or relevant information retrieved from videos recorded by cameras that represent visual sensors in this case. Metadata refers to any relevant information such as the number of moving objects in a given room, the area the moving objects occupy in the room of question, or the type of the moving objects (human or animal). The metadata inputs are also further processed to formulate appropriate independent inputs that should be fed to the fusion model. These visual independent inputs can be either mass functions or likelihood functions if using a Dempster-Shafer fusion model or a Bayesian model respectively. These inputs are labeled as $\square_1, \square_2, \square_3 \ldots, \square_{\square'}$ on figure 2.

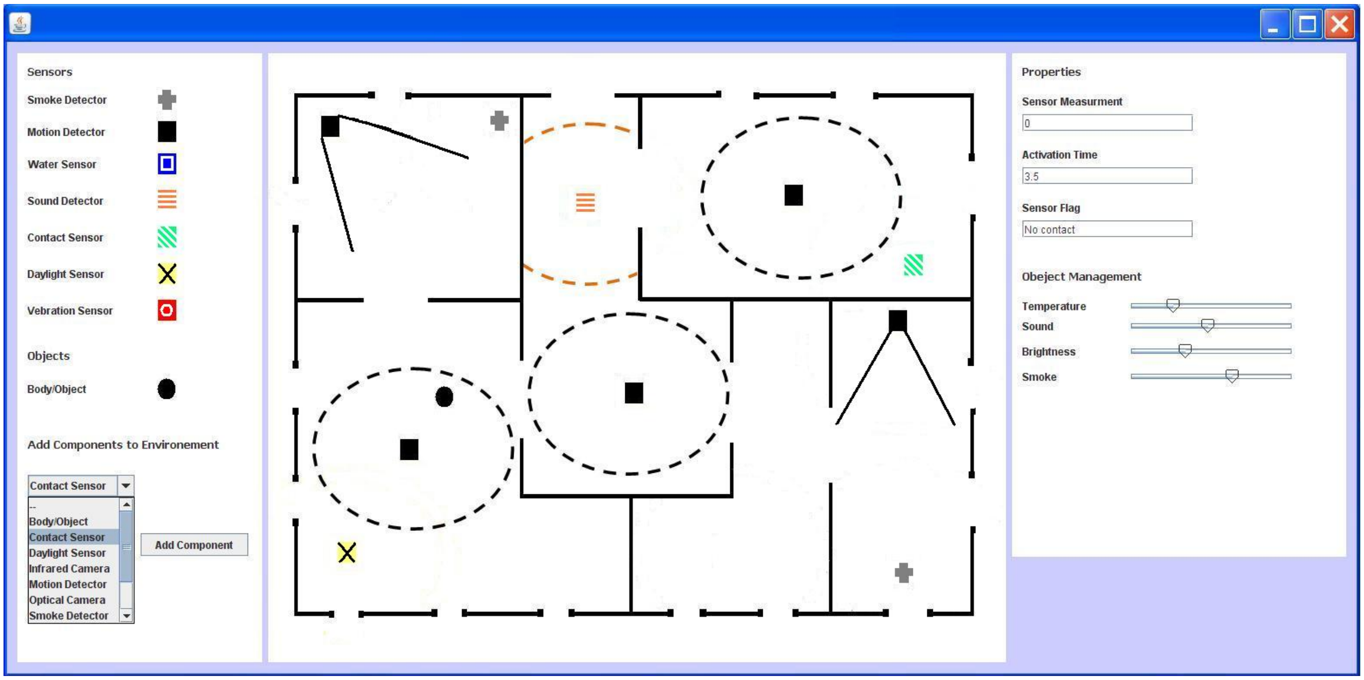


Figure1: Smart Environment Simulation Tool

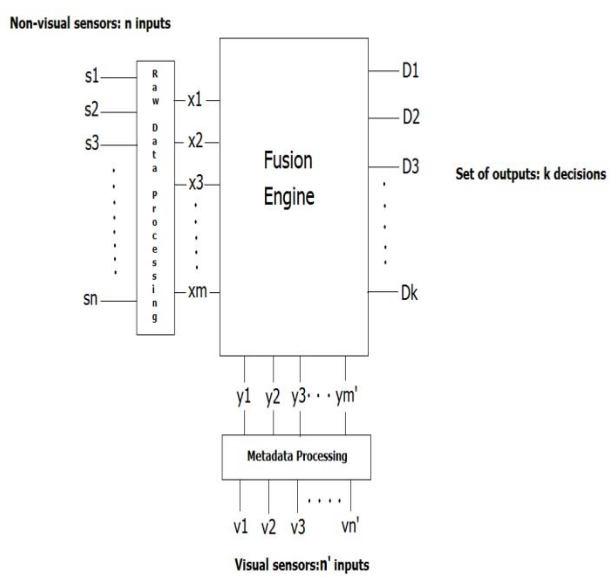


Figure 2: Data fusion engine for Decision Making in Smart Environments

The final outputs, □$_1$,□$_2$,…□$_□$, from the data fusion model shown in Figure 2 are in the form of relevant and accurate decisions. These decisions can effectively and accurately serve different environments such as surveillance, energy saving, or rescue operations. In fact, this illustrates the most important gain of data fusion, which is considering different sets of sensors to provide decisions of better quality that can safely be used in further critical and sensitive decisions.

In our study of multi-sensor data fusion, we implement a simulation tool shown in Figure 1 above. This tool helps us construct a virtual smart environment for the purposes of testing several hypotheses and building case scenarios. The smart environment has basically different types of sensors such as: motion detector, smoke detector, daylight sensor, and other types of sensors. In addition to sensors, there are objects that can be moving to generate case scenarios where motion is a factor to be considered. Emergency cases such as fire or flood can be studied using the implemented simulation tool. This tool is implemented using JAVA and it facilitates the study of multiple scenarios because the user can chose any type of sensors implemented in the tool as well as manage the environment's state such as increasing the temperature (fire case) or adding moving objects or water (flood scenario). A specific set of attributes must be defined for each sensor. These may include range, angle, sensitivity, and direction. Every sensor has a detection area and detection occurs when the coverage area and attributes of a given object overlap with the detection range and sensitivities of a given sensor. The simulation tool is our main data generator where sensors' flags and data are fed to the fusion engine where decision making process takes place.

In case of studying moving object scenario, it is preferable to place specific sensors (motion detector, sound detector, contact sensor…) as well as the moving object in our virtual environment presented by the simulation tool. As soon as the object intersects with the sensors' coverage area, we read the inputs from all the sensors that detect the

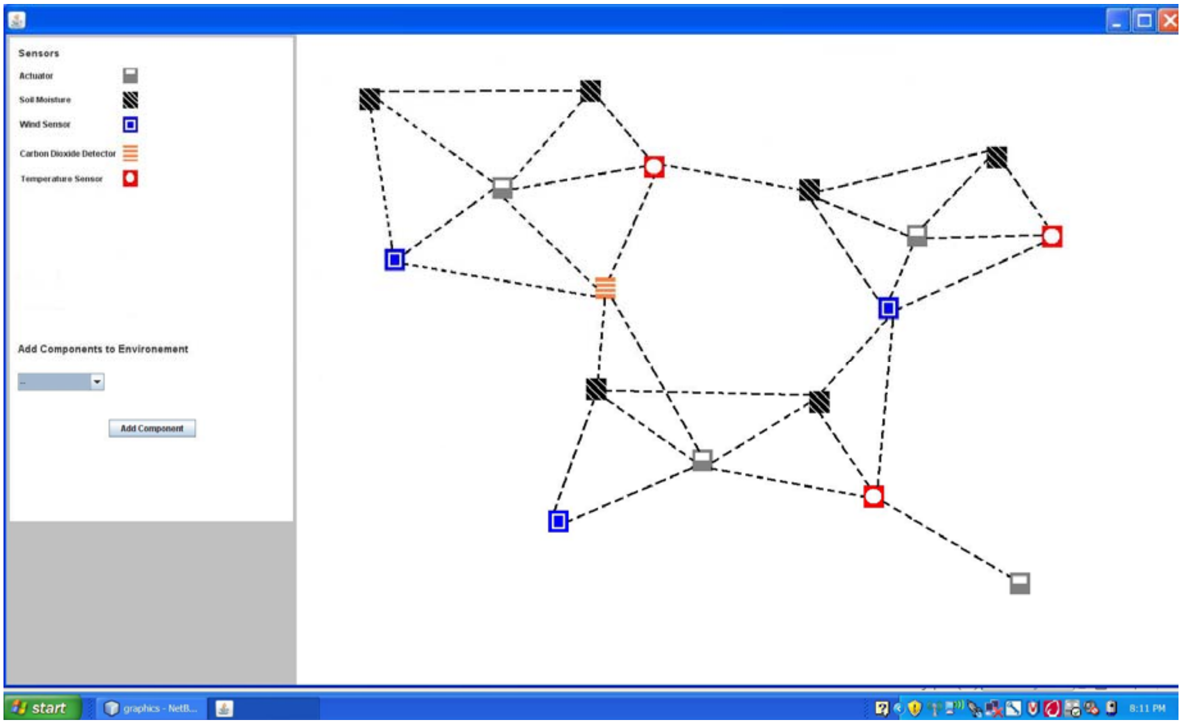


Figure 2: Smart Environment Design Tool with a Wireless Sensor Network for Irrigation

event and those inputs are presented by $\square_1, \square_2, \ldots, \square_{\square}$ in our fusion model. The independent inputs provided by simulation tool (non-visual sensors inputs) are then fused along with the independent inputs from visual sensors $\square_1$, $\square_2, \ldots$, $\square_{\square}$, using the data fusion model. The result of the fusion engine is a list of useful and relevant decisions $\square_1$, $\square_2$, …, $\square_{\square}$. These decisions could be related to what alarms need to be triggered, how many security guards need to be assigned to handle the event and such.

Generally, the smart environment concept can be used in many applications that involve detecting a current context in the environment and determining what actions should be taken based on this context information. There are many applications where this concept can be used; however, the main issue that manifests is usually that of where sensors should be placed and how many of them are needed. As a result, system optimization can be a very critical step that should take into consideration many other factors and priorities.

Design Optimization in Smart Environments

In this section we discuss a key point that should be taken into consideration when building a smart environment: optimized design (placement) of sensor and networked nodes. Sensor placement is very crucial because it influences the resource management and the type of back-end processing and exploitation that must be carried out with sensed data in distributed sensor networks [15]. The main issue here is to know where exactly these sensors need to be placed and how many sensors are needed for the optimum network performance and the cost of the system.

In outdoor applications such as agricultural irrigation, sensor placement needs to be done carefully in order to optimize the sensor resources and costs. In indoor applications as well, intelligent sensor placement facilitates the unified design and operation of sensor/ exploitation systems, and decreases the need for excessive network communication for surveillance, target location and tracking. In fact, the use of sensors should take into consideration any obstacles that might interfere with the line of vision for IR sensors. These obstacles range from buildings to trees to uneven surfaces [15].

Any approach for such an optimization should minimize the number of sensors used in the distributed network as well as decreasing the costs and optimize the amount of data that is transferred in the network. Optimized sensor placement ensures that the resulting data contains sufficient information for the data processing center to make the decisions with sufficient data. It is discussed in [16] that there exists a close resemblance between the sensor placement problem and the guard placement problem (AGP) addressed by the art gallery theorem.

Basically, the AGP problem deals with determination of the minimum number of guards required to cover the interior of an art gallery where the interior of the art gallery is presented by a polygon. Additionally, the sensor placement problem for target location is also closely related to the alarm placement problem. This problem deals with the placement of alarms on the nodes of a specific graph such that a single fault in the system (corresponding to a single faulty node in the graph) can be diagnosed. Furthermore, integer linear programming approach was also used to solve the problem of sensor placement on two and three dimensional grids. However, this approach has two main drawbacks: the complexity of computations makes it not every appropriate for large problems and the sensors are expected to be perfect where they need to yield a binary yes/no detection in each case [9].

In our optimization approach, we implemented a design optimization tool as shown in Figure 3. This tool visualizes the outcome of the optimization solution and allows the user to modify the layout manually. The tool also helps the user optimize the placement of four types of sensors commonly used in the irrigation application, chosen here as an example. These sensor types are Soil moisture sensor, temperature sensor, wind sensor, and the carbon dioxide detector. The network also includes several actuators such as valves or gates that can impact the sensors' readings and with their decisions. The design tool facilitates the building of the best possible virtual RF mesh network that will help optimize the number and location of the sensors needed and the cost of the actual implementation of such network. There are many factors that should be taken into consideration when placing sensors in an agricultural field: distance as well as obstacles between sensors or nodes in our network and the type of sensors that will be used in each case scenario. Since in our study we are considering RF network, the distance between each node in our network is very critical. As Figure 3 demonstrates, even if our network is a mesh network, we notice that some sensors are not connected either because are placed far away from one another or there are obstacles that block the communication gate. The actuators are the control units of the network that facilitate the communication between the sensors in the network as well as make the decisions needed when action is needed. For example, in case a soil moisture sensor indicates that a specific zone is dry; the actuator actually controls the valve that allows water flow to that zone.

3. Smart Environment Applications

There are many applications where an optimized smart environment can improve the effectiveness and efficiency in a particular context. In this section, we identify several such applications and discuss the benefits a smart environment could provide with the tools we have covered in the previous section.

Energy Efficiency

Energy efficiency has become a real concern in this millennium. With the high technology and the new inventions, the need of energy augments to a stage where it is a necessity to manage the energy usage in order to prevent possible losses and costs. In buildings equipped with smart features, energy efficiency has been a significant benefit. In order to assist the energy saving process, most of smart buildings are equipped with day light sensors that are an innovative energy saving device. It detects an influx of daylight, and in turn automatically dims a fluorescent luminary, or series of luminaries [11]. Daylight sensor detects any kind of light and can be used to adjust the lighting in the room to meet the needs of the room occupants. In addition, occupancy sensors can be used in smart homes/buildings to control the usage of energy for both lighting and heating/cooling areas. For example, the rooms are illuminated or heated/cooled unless they are occupied; the occupancy sensor detects the human presence and automatically turns on the room's light or heating. For such a system to work efficiently, a multitude of sensors need to be placed in various locations in a building and the inputs from multiple sensors are used make decisions at various actuator points. The interactions and relationship between these inputs and decisions are modeled using the data fusion model in Figure 2 and the network is optimized using the tools in previous section.

Surveillance

Since smart environments are always equipped with multiple sensors and processing mechanisms, they benefit significantly most of the surveillance applications. It is obvious that when fusing data from different sources, one gets a better idea about the environment and that will enhance the decision making in case of events of interest such as a robbery or fire. In our previous work on data fusion in smart environments, we presented methods for integrating surveillance camera data with data from different sensors types' in order to detect the occupancy (as well as the number of people, suspicious presence etc.) of rooms or buildings. The control of the environment, that smart environments provide, can also be very beneficial in critical situations such as fire. In such cases, the information gathered from sensors provide accurate data about where the fire spots are, which can facilitate the evacuation operation.

Irrigation

Irrigation is another important application area where smart systems have improved water and money savings. Dukes explains that irrigation controllers that have been in use since the early 2000's are smart controllers that effectively reduce outdoor water use through monitoring site conditions such as soil moisture, plant type, or wind, and irrigating based on those parameters [9]. In addition, these smart irrigation controllers receive feedback from the irrigated system and schedule the irrigation duration or frequency accordingly. An example that explains how water and money can be saved would be increasing watering the soil in hot or dry seasons and reducing it during cooler seasons. Generally, there are two types of smart controllers: climatologically-based controllers and soil moisture-based controllers [9].

Climatologically-based controllers are also known as evapotranspiration, or ET, controllers. In fact, ET is the process of transpiration by plants combined with evaporation that occurs from plant and soil surfaces. In general, three types of ET controllers are distinguished: signal based, historical ET, and on-site weather measurement. Signal based ET controllers receive meteorological data from public sources or weather stations. An ET value is then calculated for a hypothetical grass surface for that site and sent to the surrounding controllers via wireless communication. The ET controller adjusts the irrigation times or days according to the climate throughout the year. The on-site weather measurement approach, on the other hand, makes use of measured weather data at the controller to calculate ET in a continuous manner and adjust the irrigation times according to the weather conditions [9].

Alternatively, soil moisture sensor controllers make use of two control strategies: "bypass" and "on-demand". The "bypass" strategy is widely used in small sites especially residential sites. In fact, the bypass soil moisture sensor controller includes a soil moisture threshold adjustment (dry to wet) that can be used to increase or decrease the sensitivity or the point at which irrigation is needed. If the current soil moisture content exceeds the threshold, this controller delays the timed irrigation. Usually, only one soil moisture sensor is used, which requires the sensor to be placed in the driest area and adjust the run times for other areas to avoid over-watering. The on-demand soil moisture sensor controller, however, starts the irrigation at a pre-programmed low soil moisture threshold and terminates irrigation at a high threshold. This type of controllers is often used in sites that involve many irrigation zones; therefore, it initiates and terminates irrigation run times in contrast to the bypass configuration that only allows irrigation events [9].

The smart environment design and simulation tools and the theoretical models we have discussed in this paper can greatly benefit irrigation applications. In most situations, at design time, the placement of the zones and the sensors within zones is an open question. Optimized sensor network design ensures the proper network operation and that the goals in the application, such as maintaining the soil moisture level, is achieved. Data fusion methods help utilize integrated and efficient processing of sensor information and better decisions to be made as a result.

4. Conclusion and Future Work

In this paper, we have presented two prominent aspects of smart environments, namely optimized network design and data fusion within the context of several applications. We have presented a model for Dempster-Shafer data fusion technique to be effectively used in a smart environment with a heterogeneous, inter-dependent set of sensors as a data fusion technique. Additionally, we have shown that data fusion can effectively contribute to the decision making process through correcting some of information processing issues. Furthermore, we have presented design and simulation tools that may be used for the construction of smart environments with optimized sensors placement and verification of "smart features". As a next step, we are planning to implement an actual smart environment with different types of sensors and test our theoretical model (Dempster-Shafer) to fuse data in a heterogeneous network with real-time measurement data.

References

[7] "Smart Environment". Wikipedia. Feb 21, 2010 <http://en.wikipedia.org/wiki/Smart_environment>.

[8] Klein, Lehner. "Representing Data Quality in Sensor Data Streaming Environments". Feb 21, 2010 <http://portal.acm.org/citation.cfm?id=1577840.1577845 >.

[9] Dukes, Michael. "Smart Irrigation Controllers: What Makes an Irrigation Controller Smart?". University of Florida. Feb 21, 2010 <http://edis.ifas.ufl.edu/pdffiles/AE/AE44200.pdf>.

[10] Lewis, F.L. "Wireless Sensor Networks". Automation and Robotics Research Institute. <http://arri.uta.edu/acs/networks/WirelessSensorNetChap04.pdf>.

[11] Deng, Han, Mishra. "Security Support for In Network Processing in Wireless Sensor Networks". <http://www.cs.colorado.edu/~mishras/research/papers/sasn03.pdf>.

[12] Saraogi, Mayank. "Security in Wirless Sensor Networks". University of Tennessee. Feb 21, 2010 <http://www.cs.utk.edu/~saraogi/594paper.pdf>.

[13] H.B, Mitchell, *Multi-Sensor Data Fusion*. Berlin: Springer, 2007.

[14] Waltz, Edward, James, L. ,*Multisensor Data Fusion*. Norwood: Artech House, Inc, 1990.

[15] Dhillon, Chakrabarty. "Sensor Placement for Effective Coverage and Surveillance in Distributed Sensor Networks". Duke University. Feb 21, 2010

[16] Chakrabarty, Iyengar, Qi. "Grid Coverage for Surveillance and Target Location in Distributed Sensor Networks". <http://csc.lsu.edu/sensor_web/final%20papers/Gridcoverageforsurveillance.pdf >.

[17] "Smart Environments". Corelite. Feb 21, 2010 <http://www.cooperlighting.com/specfiles/brochures/SmartEnvironments-FINAL.pdf>.

[18] "BAYESIAN DATA FUSION FOR SMART ENVIRONMENTS with HETEROGENOUS SENSORS" Soukaina Messaoudi, Kamilia Messaoudi and Serhan Dagtas, Proceedings of CCSC: Mid-South 2010 8th Annual Consortium for Computing Sciences in Colleges Mid-South Conference,
Searcy, AR March 26-27, 2010

LIST OF ACRONYMS

2D	Two Dimensional
3D	Three Dimensional
4-D	Four Dimensional
AGP	Art Gallery Problem
API	Application Programming Interface
CAVE	Cave Automatic Virtual Environment
CAVELib	Cave Automatic Virtual Environment Library (of Utility Programs)
CIELAB	Commission Internationale de L'éclairage Color Model (LAB)
DAE	Digital Asset Exchange (file format)
DMOS	Difference Mean Opinion Score
ET	Evapotranspiration
FEMA	Federal Emergency Management Agency
FR	Full Reference
GAFFE	Gaze-Attention Fixation Finding Engine
GB	Gigabyte
GIS	Geographic Information System
GL	Graphical Language (programming)
GUI	Graphic User Interface
IEEE	Institute for Electrical and Electronic Engineers
IPTV	Internet Protocol Television
IQ	Information Quality
IRB	Institutional Review Board
iRC	Internet Relay Chat
IVC	Image and Vision Computing
JOGL	Java Open Graphics Library
JPEG	Joint Photographic Experts Group
LAB	Color model where L represents Lightness, and “A” and “B” represent two color dimensions
LAR	Local Adaptive Resolution
LIVE	Laboratory for Image & Video Engineering, University of Texas at Austin
LMS	Long, Medium, and Short Wavelength (human eye color space)
MATLAB	Matrix Laboratory
MOS	Mean Opinion Score
MSE	Mean Squared Error
NASA	National Aeronautics and Space Administration
NR	No Reference
OpenCV	Open Source Computer Vision
PDA	Personal Digital Assistant
QoE	Quality of Experience
RBF	Radial Basis Function
RF	Radio Frequency
RGB	Red, Green, and Blue (additive color model)
RR	Reduced Reference
RSS	Really Simple Syndication
SG&A	Computer Graphics & Applications

SNA	Social Network Analysis
SPIE	International Society for Optics and Photonics
SQL	Standard Query Language
S-SIM	Saliency-Based Structural Similarity Index
S-VIF	Saliency-Based Visual Information Fidelity
UALR	University of Arkansas at Little Rock
USGS	United State Geographical Survey
UTI	Unstructured Textual Information
VIF	Visual Information Fidelity
VQA	Video Quality Assessment
WEKA	Wikato Environment for Knowledge Analysis
WPAFB	Wright-Patterson Air Force Base