Final Report on Grant DAAH04-94-G-0284, Image Descriptions for Browsing and Retrieval

Carlo Tomasi and Leonidas J. Guibas
Stanford University

1 Report Structure

In the next section, we outline the problem studied under this grant. In section 3 we describe our accomplishments. Section 4 shows the papers published, and 5 lists students who have been funded by this grant.

2 Problem Studied

The literature on image retrieval is growing, with several efforts in both academia [9, 7, 13, 21, 18, 26, 15, 22] and industry [6, 25, 8]. The main thrust of our work is the definition of basic image representations that are most appropriate for image search. With the aim of a unified treatment, we have developed the notion of a *signature* to summarize image appearance. Signatures can represent the color, shape, or texture content of an image. They are more flexible than feature vectors and histograms, as they imply no fixed number or ordering of feature primitives, as in vectors, nor fixed-pitch quantization of feature values, as in histograms. Color, shape, and texture signatures are described in sections 3.1, 3.4, and 3.5.

By using a single representation format for the three different modalities considered in our work, that is, color, shape, and texture, we have made our retrieval mechanisms essentially uniform across modalities. This has led not only to efficiency and simplicity, but also to conceptual consistency.

The other main ingredient of a retrieval system, besides signatures, is a perceptually meaningful measure of similarity between two images. We have defined such a measure based on what we call the "Earth Mover's Distance" (section 3.2). With these two ingredients, the pictures in a database can be organized so as to keep similar images close to each other. In this context, we have developed efficient data structures for sublinear nearest-neighbor retrieval. In addition, a similarity metric between images leads to methods for laying out either all the images in the database, or a sample thereof, or a small number of mutually related images, and for displaying these images in an intuitive way for the user. The mathematical tool we used for the creation of this layout is multi-dimensional scaling (MDS).

In shape-based retrieval, we have used shape information in the presence of occlusions to retrieve drawings from various colelctions, and we have developed shape indices by recording "what basic shape appears where in the image." We successfully experimented with data-bases of illustrations from geometry textbooks, and of scanned-in Chinese characters. For this work, we extended geometric hashing techniques to make our indices invariant under a transformation group.

We have built a web-based retrieval system that allows fast retrieval from a 20,000 image database based on color signatures (section 3.6). Furthermore, we demonstrated the notion of a database navigator, in which many of the images in a database are laid out in three-dimensional space (section 3.6). The user then navigates in this space with a joystick. The main advantage of this new interaction paradigm is that the content of the database is conveyed to the user all at once, rather than piecemeal, as in the more standard query/response protocol. A global view lets the user form a mental picture of the database, just as one forms a mental picture of the contents of, say, a bookstore by browsing in it for some time. If the images are arranged in a coherent fashion, consistent with our similarity metric, the ordering rationale is easily learned by the user without being explicitly identified. At a more local level, again thanks to our metric, the small number of images returned in response to a query in the more traditional query/response operating mode can be displayed so as to emphasize similarities and differences among the images.

A retrieval system we developed for police mugshots (section 3.7) demonstrates the usefulness of signatures, EMD, and our navigation tools for a real-world application.

In the following section, we outline the main achievements of our work.

3 Summary of Results

3.1 Color Signatures

The color information of each image is reduced to a compact representation that we call the *signature* of the image. In general a signature contains a varying number of points in a Euclidean space where a weight is attached to each point. In the case of color images, the points represent clusters of similar colors in CIE-LAB space, and the weight of a point is the fraction of the image area with that color. The signatures thus obtained are compact: the color distribution of an entire image is summarized by a handful of points, typically eight to twelve. Since signatures represent distributions in the CIE-LAB color space, they are perceptually significant, in that Euclidean distances between points are strongly correlated with perceptual differences. Because of clustering, small variations in the colors of an image have little effect on signatures, thereby providing a moderate degree of invariance to changes of viewpoint and lighting. Finally, signatures are simple and flexible abstractions. In fact, the cloud of weighted points that makes up a color signature lives in the low-dimensional space of colors. Furthermore, just as objects and concepts are described in English by sentences with a variable number of words, so images are summarized by a variable number of colors in a signature. The ordering of colors in not meaningful, and is therefore not used. The relative importance of the various colors is explicitly represented by the weight of each signature component, and is therefore immune from the quantization problems inherent in color histograms.

3.2 The Earth Mover's Distance

We define the distance between two signatures to be the minimum amount of 'work' needed to transform one signature into the other. The work needed to move a point, or a fraction of a point, to a new location is the portion of the weight being moved, multiplied by the Euclidean distance between the old and the new locations. When changing one signature to another, the work is the sum of the work done by moving the weights of the individual points of the source signature to those of the destination signature. We allow the weight of a single source signature point to be partitioned among several destination signature points, and vice versa. The distance between the source and destination signatures is then defined to be the minimum amount of work necessary to thus move the weight of the source to that of the destination signature. We call this distance function the *earth mover's distance*.

Computing the earth mover's distance can be formulated as a linear programming (LP) problem [16]. Given the compact nature of color signatures, this LP problem is relatively small. Still, since computing this distance is the main operation in our image retrieval systems, we are devoting considerable efforts to making this solution as fast as possible. The distance between two images is computed in a few milliseconds. We have developed bounds that can be used both to exclude from consideration images that are too distant from the query and to abort computation of a distance once it is certain to exceed a certain value.

3.3 Color Metric Comparisons

Multidimensional distributions are often used in computer vision to describe and summarize the color content of an image. Given two distributions of colors, it is often useful to define a quantitative measure of their dissimilarity, with the intent of approximating perceptual dissimilarity as well as possible. This is particularly important in image retrieval applications, but has fundamental implications also for the understanding of color perception. Defining a distance between two distributions requires first a notion of distance between the basic features that are aggregated into the distributions. We call this distance the *ground distance*. For instance, in the case of color, the ground distance measure dissimilarity between individual colors. Fortunately, color ground distance has been carefully studied in the literature of psychophysics, and has led to measures like the CIE-Lab color space [27].

Given a ground distance, several measures have been proposed in the literature for the perceptual dissimilarity of color distributions and distributions in general. In this research, we surveyed some of these measures, and compared them with the Earth Mover's Distance (EMD), a metric we proposed in [17]. The EMD can be applied both to histograms and to so-called signatures, which are more flexible representations of distributions. We showed that the combination of EMD with signatures works best for image retrieval.

3.4 Shape-Based Illustration Indexing and Retrieval

We have developed a general set of ideas for indexing computer-generated technical illustrations based on the shapes present in them, so that they can be efficiently retrieved later using as the key other 'similar-looking' illustrations (either pre-existing, or interactively drawn by the user). We have restricted out attention to the domain of computer-generated technical illustrations for now, where shape information is both precisely available and the main way in which pictorial meaning is conveyed. After we have techniques that can operate successfully in this domain, we plan to port them to other kinds of pictorial or image data as well by applying shape extraction techniques from computer vision.

We proceed as follows: given an illustration P, we compute a compact index $\iota(P)$ which records the principal shapes present in P and their location/orientation/size. Then given a collection of illustrations, we compute a data-structure for recording their indices so that later queries can be answered efficiently. At retrieval time we are given another illustration Q; we compute $\iota(Q)$ and then search the data base for illustrations P_i whose index $\iota(P_i)$ is 'similar' to $\iota(Q)$.

An illustration P for us is a collection of instanced graphics primitives (lines or polylines, circular arcs, Bézier cubic or B-spline arcs, marks, etc.), as is almost universally the case with the illustrators in common use today (e.g., Adobe Illustrator, Aldus Freehand, Xfig, etc.). We start with a a collection of *basic shapes* which may be built-in, or user-definable. In the index $\iota(P)$ of an illustration P we record 'which basic shapes appear where.' In other words, for each basic shape, we record in the index the translation, rotation, and scale transformations which cause this basic shape to match well some of the shapes present in P, according to the Hausdorff distance [3]. Thus we can think of the index as a list of 'colored' points in $\mathcal{R}^4$, where the four coordinates are the four parameters defining the transformation, and the color is the label of the basic shape involved. (We actually store the logarithm of the scale parameter, so as to make variations in scale correspond to point translations in $\mathcal{R}^4$, just like for translations and rotations).

When a query illustration comes in, we compute its index $\iota(Q)$ in the same way. At the moment we match $\iota(Q)$ with the index of every illustration in the data base, by computing in $\mathcal{R}^4$ the *colored one-way Hausdorff distance under translation* between the two point sets representing the indices;

a fuller explanation of the matching mechanism is given in [5]. We are optimistic that in the future we will be able to attain sublinear query-time algorithms (algorithms which do not need to compare $\iota(Q)$ with every other illustration index) by using computational geometric techniques on the set of indices — essentially by clustering illustrations whose indices have a small 'distance' from each other.

A library of approximately two-hundred illustrations from a geometry textbook was indexed using this scheme and then used for retrieval experiments. An interactive interface was provided for specifying the data base to be searched and the query illustration, for setting various parameters regarding the match, and for displaying the best matches found in the data base. Details and examples are provided in [5].

3.5 Texture Metrics

Similarity measures between textures are important for image understanding applications such as content-based image retrieval, texture segmentation, and texture classification. In order to be useful, it is important that these similarity measures correspond to human texture perception. In addition, in image retrieval it is often crucial that the similarity distances be *metric*, so that efficient data structures and search algorithms [2, 4] can be used.

In this research we defined a class of texture metrics based on texture features close to the model of simple cells in the primary visual cortex[11]. For the distance between texture feature histograms we used the Earth Mover's Distance (EMD), an effective and efficient measure of histogram differences [17]. We evaluated our metrics both quantitatively, by examining the actual distances between different textures, and qualitatively, by using multidimensional scaling techniques [24] to find what are the texture properties that affect our metrics the most, and to "visualize" the metrics. We obtained similar results to those found by psychophysical experiments [23, 14], thereby confirming our claim that our texture metrics correspond to human perception.

3.6 Database Navigation

The user of an image retrieval system would typically like to specify queries in semantic terms (e.g. "children playing in a park"). Unfortunately, the state-of-art in computer vision does not yet allow for such queries. Instead, systems use simpler syntactic image features such as color, texture and shape [6, 1, 7, 12, 13], in the hope that these correlate well with semantic features. This discrepancy between syntactic and semantic queries causes a basic problem with the traditional query/response style of interaction. An overly generic query yields a large jumble of images, which are hard to examine, while an excessively specific query may cause many good images to be overlooked by the system. This is the traditional trade-off between good precision (few false positives) and good recall (few false negatives). Striving for both good precision and good recall may pose an excessive burden on the definition of a "correct" measure of image similarity. While most image retrieval systems, including the ones above, recognize this and allow for an iterative refinement of queries, the number of images returned for each query is usually kept low so that the user can examine them one at a time.

In contrast, we suggested that with an appropriate display technique, which is the main point of this research, many more images can be returned without overloading the user's attention. Specifically, if images can be arranged on the screen so as to reflect similarities and differences between their color distributions, the initial queries can be very generic, and return a large number of images. The consequent low initial precision is an advantage rather than a weakness. In fact, the user can see large portions of the database at a glance, and form a global mental model of what is in it. Rather than following a thin path of images from query to query, as in the traditional approach, the user now *zooms in* to the images of interest. Precision is added incrementally in subsequent query refinements, and fewer and fewer images are displayed as the desired images are approached.

In our system, we use the distributions of colors in images as our retrieval features. These have been shown [22, 6, 20, 1, 7, 12, 13] to be useful retrieval cues. When a (usually vague) query is specified or drawn by the user, we locate and display a large number of neighboring images in the database. Since queries in our system are image-like, neighborhood can be defined in terms of the distance between images. The resulting images are then used for more focused queries that return fewer and fewer images. At every step, query results are embedded in two-dimensional space by using *multi-dimensional scaling (MDS)* [19, 10], by which we place picture thumbnails on the screen so that screen distances reflect as closely as possible the distances between the images. While more traditional displays list images in order of similarity to the query, thereby representing n distances if n images are returned, our display conveys information about all the $\binom{n}{2}$ distances between images. This display makes it easy for the user to grasp the entire set of returned images at a glance, understand how the query actually performed, and decide where to go next. In fact, such geometric embeddings allow the user to perceive the dominant axes of variation in the displayed image group. When the user selects a region of interest on the display, a new, more specific query is automatically generated, and returns a smaller set of images. These are again displayed by a new MDS, which now reflects the new dominant axes of variation. Thus, the embeddings are *adaptive*, in the sense that they use the screen's real estate to emphasize whatever happen to be the main differences and similarities among the particular images at hand. By iterating this process, the user is able to quickly navigate to the portion of the image space of interest, typically in very few mouse clicks.

3.7 Mugshot Retrieval

A police mugshot retrieval system was developed as a feasibility test for a Canadian police department. In order to identify a suspect in a crime, witnesses are often asked to scan large collections of police mugshots. Fatigue and insufficient attention span can lead to distraction during this process. A system that lets witnesses *navigate* through the mugshot collection in a more coherent fashion can help reduce the likelihood of costly mistakes.

The witness gives general indications about the appearance of the suspect, such as age group, sex, race, and hair color. The system then displays many relevant images as small thumbnail icons on a single screen, arranging them in such a way that similar faces, in terms of the attributes of importance for the given search, appear close to each other on the screen. Because similar images are nearby, it becomes much easier for a witness to concentrate his or her attention on the part of display of interest. By selecting an "interesting" part of the display, the system produces a new display, with images that are similar to those in the interesting part. By repeating this procedure, the witness can home in to the image of the suspect in a few steps.

In order to apply our perceptual navigation algorithms, we needed a set of images and a similarity measure between them. The image set was provided by a Canadian police department. We used a very simple feature-based similarity measure between mugshots, based on simple, pre-annotated features provided by the police. We defined a distance measure for every feature, and the similarity measure between two mugshots was defined as a linear combination of these distances. The relative importance of the features was controlled by modifying the weights of each feature. The user could also "turn off" features which should not participate in the computation of the similarity measure.

4 Publications

The following publications have been produced during this grant.

R. Manduchi and C. Tomasi. Distinctiveness maps for image matching. In *10th International Conference on Image Analysis and Processing (ICIAP)*, Venice, Italy, September 1999, pages 26–31.

Y. Rubner and C. Tomasi. Texture-based image retrieval without segmentation. In *Seventh International Conference on Computer Vision (ICCV)*, Kerkyra, Greece, September 1999, pages

1018–1024.

M. A. Ruzon and C. Tomasi. Corner detection in textured color images. In *Seventh International Conference on Computer Vision (ICCV)*, Kerkyra, Greece, September 1999, pages 1039–1045.

J. Puzicha, Y. Rubner, C. Tomasi, and J. M. Buhmann. Empirical evaluation of dissimilarity measures for color and texture. In *Seventh International Conference on Computer Vision (ICCV)*, Kerkyra, Greece, September 1999, pages 1165–1172.

M. A. Ruzon and C. Tomasi. Color edge detection with the compass operator. In *Proceedings of the IEEE Computer Science Conference on Computer Vision and Pattern Recognition (CVPR)*, pages II–160–166, Fort Collins, CO, June 1999.

C. Tomasi and L. J. Guibas. Retrieving Color, Patterns, Texture, and Faces. *Proceedings of the DARPA Image Understanding Workshop*, Monterey, CA, November 1998, pp. 107–111.

Y. Rubner and C. Tomasi. Comparing the EMD to Other Dissimilarity Measures for Color Images. *Proceedings of the DARPA Image Understanding Workshop*, Monterey, CA, November 1998, pp. 331–339.

Y. Rubner and C. Tomasi. Texture Metrics. *IEEE International Conference on Systems, Man, and Cybernetics*, San Diego, CA, October 1998, pages 4601–4607.

Y. Rubner, C. Tomasi, and L. J. Guibas. The Earth Mover's Distance as a Metric for Image Retrieval. *Technical Report STAN-CS-TN-98-86*, Computer Science Department, Stanford University, September 1998.

Y. Rubner, C. Tomasi and L. J. Guibas. Adaptive Color-Image Embeddings for Database Navigation. *Proceedings of the Third Asian Conference on Computer Vision*, Hong Kong, pages 104–111, January 1998.

C. Tomasi and R. Manduchi. Bilateral Filtering for Gray and Color Images. *Proceedings of the Sixth International Conference on Computer Vision (ICCV)*, Bombay, India, pp. 839–846, January 1998.

Y. Rubner and C. Tomasi and L. Guibas. A Metric for Distributions with Applications to Image Databases. *Proceedings of the Sixth International Conference on Computer Vision (ICCV)*, Bombay, India, pp. 59–66, January 1998.

L. J. Guibas and C. Tomasi. Image Browsing and Retrieval Research at Stanford. *Proceedings of the DARPA Image Understanding Workshop*, New Orleans, LA, May 1997, pp. 495–500.

Y. Rubner, L. J. Guibas, and C. Tomasi. The Earth Mover's Distance, Multidimensional Scaling, and Color-Based Image Retrieval. *Proceedings of the DARPA Image Understanding Workshop*, New Orleans, LA, May 1997, pp. 661–668.

L. J. Guibas and C. Tomasi. Image Retrieval and Robot Vision Research at Stanford. *Proceedings of the ARPA Image Understanding Workshop*, Palm Springs, CA, March 1996, pp. 101–108.

Y. Rubner and C. Tomasi Coalescing Texture Descriptors. *Proceedings of the ARPA Image Understanding Workshop*, Palm Springs, CA, March 1996, pp. 927–935.

L. Guibas and B. Rogoff and C. Tomasi. Fixed-Window Image Descriptors for Image Retrieval. *Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases*, pages 2420–2431. San Josè, CA, February 1995.

C. Tomasi and L. Guibas. Image Descriptions for Browsing and Retrieval. *Proceedings of the ARPA Image Understanding Workshop*, Monterey, CA, November 1994, pages 165–168.

5 Scientific Personnel

Principal investigators: Carlo Tomasi and Leonidas J. Guibas.

The following students have worked at various stages on this project: Byung-Hyun Chung, Scott Cohen, Aristides Gionis, David Hoffman, Eugene Jhong, Brian Rogoff, Joseph Rubner, Mark Ruzon.

References

[1] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C. Shu. Virage image search engine: an open framework for image management. In *SPIE Conference on Storage and Retrieval for Image and Video Databases IV*, volume 2670, pages 76–87, March 1996.

[2] T. Bozkaya and M. Ozsoyoglu. Distance-based indexing for high-dimensional metric spaces. *SIGMOD Record (ACM Special Interest Group on Management of Data)*, 26(2):357–368, May 1997.

[3] L. P. Chew, M. T. Goodrich, D. P. Huttenlocher, K. Kedem, J. M. Kleinberg, and D. Kravets. Geometric pattern matching under Euclidean motion. In *Proceedings of the Fifth Canadian Conference on Computational Geometry*, pages 151–156, 1993.

[4] Kenneth L. Clarkson. Nearest neighbor queries in metric spaces. In *ACM Symposium on the Theory of Computing*, pages 609–617, 1997.

[5] S. D. Cohen and L. J. Guibas. Shape-based illustration indexing and retrieval: some first steps. In *Proceedings of the ARPA Image Understanding Workshop*, pages 1209–1212, Palm Springs, CA, 1996.

[6] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*, 3:231–262, 1994.

[7] D. Forsyth, J. Malik, M. Fleck, H. Greenspan, and T. Leung. Finding pictures of objects in large collections of images. In *International Workshop on Object Recognition for Computer Vision*, Cambridge, UK, April 1996.

[8] S. I. Gallant and M. F. Johnston. Image retrieval using image context vectors: first results. In *IS&T/SPIE Symposium on Electronic Imaging: Science and Technology*, volume 2420, pages 82–94, San Josè, CA, February 1995.

[9] L. J. Guibas and C. Tomasi. Image retrieval and robot vision research at Stanford. In *Proceedings of the ARPA Image Understanding Workshop*, pages 101–108, Palm Springs, CA, 1996.

[10] J. B. Kruskal. Multi-dimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964.

[11] T. S. Lee. Image representation using 2d Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971, 1996.

[12] Greg Pass and Ramin Zabih. Histogram refinement for content-based image retrieval. In *IEEE Workshop on Applications of Computer Vision*, Sarasota, FL, 1996.

[13] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, June 1996.

[14] R. A. Rao and G. L. Lohse. Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, 55:218–233, 1993.

[15] S. Ravela, R. Manmatha, and E. M. Riseman. Image retrieval using scale space matching. In B. Buxton and R. Cipolla, editors, *Computer Vision - ECCV96*, pages 273–282, Cambridge, UK, April 1996. Springer-Verlag.

[16] Y. Rubner, L. J. Guibas, and C. Tomasi. The earth mover's distance, multidimensional scaling, and color-based image retrieval. In *Proceedings of the DARPA Image Understanding Workshop*, pages 661–668, New Orleans, LA, 1997.

[17] Y. Rubner, C. Tomasi, and L. Guibas. A metric for distributions with applications to image databases. In *Proceedings of the Sixth International Conference on Computer Vision (ICCV)*, pages 59–66, Bombay, India, January 1998.

[18] S. Santini and R. Jain. Similarity queries in image databases. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR96)*, pages 646–651, San Francisco, CA, 1996.

[19] R. N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function, i and ii. *Psychometrika*, 27:125–140,219–246, 1962.

[20] M. Stricker and M. Orengo. Similarity of color images. In *SPIE Conference on Storage and Retrieval for Image and Video Databases III*, volume 2420, pages 381–392, February 1995.

[21] M. A. Stricker. Color indexing with weak spatial constraints. In I. K. Sethi and R. C. Jain, editors, *Photonics West '96 — Storage and Retrieval for Still Image and Video Databases IV*, pages 2670–04, San Josè, CA, February 1996.

[22] M. J. Swain and D. H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

[23] H. Tamura, T. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man and Cybernetics*, 8:460–473, June 1978.

[24] W. S. Torgerson. *Theory and Methods of Scaling.* John Wiley and Sons, New York, NY, 1958.

[25] `http://www.virage.com/`, 1997.

[26] X. Wan and C. J. Kuo. Image retrieval with multiresolution color space quantization. In *Proceedings of the SPIE*, volume 2898, pages 148–159, 1996.

[27] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae.* John Wiley and Sons, New York, NY, 1982.

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE May 2000	3. REPORT TYPE AND DATES COVERED Final Report

4. TITLE AND SUBTITLE
Image Descriptions for Browsing and Retrieval

5. FUNDING NUMBERS
DAAH04-94-G-0284

6. AUTHOR(S)
Carlo Tomasi

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Stanford University
Stanford, CA 94305

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 33583.5-MA

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The main thrust of our work under this grant has been the definition of basic image representations that are most appropriate for image retrieval. With the aim of a unified treatment, we have developed the notion of a signature to represent the color, shape, or texture content of an image. We have proposed a new, perceptually motivated metric for signatures, the Earth-Mover's Distance (EMD), and designed sound and efficient algorithms for its computation. Based on these foundational elements we have developed the notion of database navigation as a novel, effective paradigm for interaction with a large database of images. These concepts were the basis for a solid list of publications. We have demonstrated all our ideas with a series of software systems for the retrieval of images from large reporitories based on color, texture, shape, and facial features. Eight PhD students have been funded by this project, and three of them have completed their PhD degrees at Stanford with theses on various aspects of image retrieval.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE