IMAGE SEQUENCE ENHANCEMENTS BASED ON
THE NORMAL COMPONENT OF IMAGE MOTION

Larry S. Davis
Ku-chen Xie
Azriel Rosenfeld

Computer Vision Laboratory
Computer Science Center
University of Maryland
College Park, MD 20742

COMPUTER SCIENCE
TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited
IMAGE SEQUENCE ENHANCEMENTS BASED ON THE NORMAL COMPONENT OF IMAGE MOTION

Larry S. Davis  
Hu-chen Xie  
Azriel Rosenfeld

Computer Vision Laboratory  
Computer Science Center  
University of Maryland  
College Park, MD 20742

ABSTRACT

Huang and Hsu [1] describe an image sequence enhancement algorithm based on computing motion vectors between successive frames and using these vectors to determine the correspondence between pixels for frame averaging. In this note, we demonstrate that it may be sufficient to use only the components of the motion vectors in the gradient direction (called the normal components) to perform the enhancement.

The support of the Defense Advanced Research Projects Agency and the U.S. Army Night Vision Laboratory under Contract DAAG-53-76C-0138 (DARPA Order 3206) is gratefully acknowledged, as is the help of Janet Salzman in preparing this paper.
1. Introduction

In a recent paper, Huang and Hsu [1] describe an enhancement algorithm for time-varying images. This algorithm achieves the type of enhancement that one would obtain from averaging multiple images of a static scene by using motion vectors computed from the sequence to establish the correspondence between pixels in successive frames.

The purpose of this note is to demonstrate how and why it is often sufficient to consider only the normal components of the motion vectors, i.e., the component in the gradient direction, to compute such enhancements. There is a computational advantage to using only the normal component, since it is directly measurable from the image sequence by combining the space and time derivatives of image intensity, while computing the actual motion vectors ordinarily requires either some spatial integration of the normal components or the application of a local correlation procedure. Furthermore, in order to compute the motion vectors one must make some assumption about the local image motion (for example, that the image motion is a two-dimensional translation); when this assumption is incorrect, the computed motion vectors will be in error.

Using the motion vectors to perform image sequence enhancement implies that it is important to establish the exact correspondence between points from frame to frame. In fact, if we assume that the objects in the image sequence have (piecewise) slowly varying
intensity, then it is sufficient to simply determine a correspondence between nearby points that belong to the same image region, since their intensities should be quite similar. The normal component can be used to establish such a correspondence. Indeed, let $P$ be a pixel, let $P^*$ be the result of displacing $P$ by the true motion vector, and let $P'$ be the result of displacing it by the normal component only. Thus $P^*$ is the true position of $P$ under the motion; but $P'$ differs from $P^*$ only by the (unknown) tangential component of the motion, and so is likely to belong to the same region as $P^*$. The concept here is analogous to that used in early work on smoothing of single images [2], where a pixel at which the gradient magnitude is high is smoothed by averaging with its neighbors in the tangential direction only.
2. Enhancement algorithm

Let $f_1$ and $f_2$ be two frames from a motion sequence, and let $m$ be the field of normal components of the motion vectors between $f_1$ and $f_2$. The picture $m$ is a vector-valued picture, where $m(x,y) = (u,v)$; $u$ is the projection of the normal component on the x-axis and $v$ is the projection of the normal component on the y-axis. We will let $u(x,y)$ and $v(x,y)$ denote the $u$ and $v$ components of $m$ at $(x,y)$.

We compute an enhanced version of $f_2$, which we will denote as $g_2$, by the following two stage process:

1) Create a "shifted" version of $f_1$, denoted $f'_1$, as follows:
   a) Initially, for all $(x,y)$ set $f'_1(x,y)=U$ (for "undefined")
   b) For all $(x,y)$ set $f'_1(x+u(x,y), y+v(x,y)) = f_1(x,y)$

2) Create the enhanced version of $f_2$ as follows:

   For all $(x,y)$
   
   If $f'_1(x,y)=U$ Then $g_2(x,y) = f_2(x,y)$
   
   Else $g_2(x,y) = \frac{f'_1(x,y) + f_2(x,y)}{2}$

In step (1) the pixels having no predecessor in the image under the motion are assigned the value "U" in the shifted version of $f_1$. At these positions we cannot compute a correspondence between pixels in $f_1$ and $f_2$, so that we use the intensities in $f_2$ to form the enhanced image. Also, it is quite possible that several points in $f_1$ map onto the same point in $f'_1$ due to errors in the normal components. Note that this could also happen if we used the motion vectors rather than only the normal components.
The above algorithm can be simply extended to compute enhancements based on sequences of more than two frames. Let \( f_1, f_2, \ldots, f_n \) be the sequence of frames for which we wish to produce an enhanced image. We proceed as follows: Let \( m_i \) be the normal component or motion vector field for the \( i \)th frame. Apply the above algorithm to frames \( f_1 \) and \( f_2 \) to produce an enhanced image of those two frames, which we will denote by \( g_{12} \). Next, apply the enhancement algorithm to \( g_{12} \) and \( f_3 \) (using the motion vector field \( m_2 \) computed from the initial sequence) to produce an enhanced version of \( f_1-f_3 \) denoted \( g_{13} \). Continuing in this manner, we can produce an enhanced version of frames \( f_1 \) to \( f_n \).

A problem with this approach is that as we compose the motion vector fields in moving through the sequence the errors in the individual motion vector fields accumulate, so that \( g_{in} \) is not a uniformly good enhancement of \( f_1 \) to \( f_n \). Instead we adopted the following "triangular" procedure which takes advantage of the intermediate computations that would have to be done to compute enhancements of subsequent sequences (e.g., \( f_2 \) to \( f_{n+1} \)) for computing the enhancement of \( f_1 \) to \( f_n \). Instead of simply regarding \( g_{in} \) as the enhanced version of \( f_1 \) to \( f_n \), we compute the median of \( g_{in}, i=1, n-1 \) and \( f_n \) as the enhancement. The images \( g_{in}, i=2, n-1 \), are intermediate results in the computation of \( g_{1,i+n-1} \) and so must be computed in any event. We found experimentally that median filtering results in better enhancements than simple averaging.
3. **Experimental results**

Figures 1 and 2 show two image sequences, each containing 8 frames. Sequence 1 contains images of a printed page which was moved in front of the camera parallel to the image plane (thus the motion was a two-dimensional motion). White noise was then added to each frame in the sequence. Sequence 2 contains images of a toy with significant detail. The camera was rotated so that the image of the toy moved from left to right in the sequence, and in this case the motion is a three-dimensional motion.

The results comparing enhancements obtained using the normal components of the motion vectors versus the actual motion vectors are displayed in Figures 3 and 4. The motion vectors and normal components were computed using a differential technique. Spatial and temporal intensity derivatives were approximated based on cubic spatial and temporal approximations to sequences of five frames, with the motion vectors associated with the third frame in each sequence of five. The motion vectors were derived from the normal components under the assumption that the motion was locally a simple image plane translation (see, for example [3,4, 5]). For comparison, enhancements based on both averaging and median filtering (see the discussion at the end of the previous section) are displayed, as well as enhancements without motion compensation. The results using the motion vectors and the normal components are quite comparable, and indicate that it may not always be necessary to incur the additional expense of determining the true motion vectors to perform enhancement.
4. Conclusions

In this paper we have demonstrated that it may not always be necessary to compute motion vectors to enhance image sequences, and that instead it may be sufficient to consider only the normal component of motion. However, we should note that this observation may depend on the specific enhancement algorithm chosen; also, since the evaluation of image enhancements is based on visual, subjective criteria we should point out that Huang [1] remarked that the enhancements are stronger when the enhanced images are themselves viewed as an image sequence. We were unable to perform such a "dynamic" comparison with normal component enhancements, but the results of such a subjective comparison would be important in deciding whether or not to use just the normal components for enhancement.
References


Figure 1. Two-dimensional motion sequence.
Figure 2. Three dimensional motion sequence.
Figure 3. Enhancements of Figure 1.
Figure 4. Enhancements of Figure 2.
IMAGE SEQUENCE ENHANCEMENTS BASED ON THE NORMAL COMPONENT OF IMAGE MOTION

Larry S. Davis
Hu-chen Xie
Azriel Rosenfeld

Computer Vision Laboratory
Computer Science Center
University of Maryland
College Park, MD 20742

U.S. Army Night Vision Lab.
Pt. Belvoir, VA 22060

Approved for public release; distribution unlimited

Huang and Liu [1] describe an image sequence enhancement algorithm based on computing motion vectors between successive frames and using these vectors to determine the correspondence between pixels for frame averaging. In this note, we demonstrate that it may be sufficient to use only the components of the motion vectors in the gradient direction (called the normal components) to perform the enhancement.