A MOTION CONFIDENCE MEASURE FROM PHASE INFORMATION

Long To, Mark Pickering, Michael Frater, John Arnold

School of Information Technology and Electrical Engineering
University College, The University of New South Wales
Australian Defence Force Academy, Canberra
{tlong,mp,mrf,jfa}@ee.adfa.edu.au

ABSTRACT

The reliance of many video processing techiniques on motion estimation requires a good motion confidence measure, in order to ensure that unreliable estimations are not accepted into a system. This paper proposes a new measure of motion confidence based on phase information from Fourier transform of a video frame. Using the measure, blocks associated with unreliable motion estimation are detected efficiently and accurately. The merit of the technique is a much improved sensitivity over circumstances where other measures show ambiguity. This improvement is demonstrated through our experimental and analytical results.

1. INTRODUCTION

Motion information is a crucial input to many applications, such as video coding, object segmentation and tracking. A motion estimation process is however subject to a number of problems, among them occlusion and a lack of aperture. These problems are not known when selecting the area for which the motion is estimated, and therefore accuracy of the resulting motion field cannot always be assumed. A wrong motion estimation would lead to incorrect formation of object masks in segmentation, inaccurate tracking results, and visual effects in coded videos. Because effects of a wrongly-estimated motion is usually difficult to correct in post-processing, it is important that any application relying on temporal information should properly address the problem of motion reliability. For example, it is considered a main issue for optimal performance of a tracking system [1].

A number of motion confidence measures have been put into use. In [2], the peak values in the phase-correlation field is offered as a goodness-of-fit measure, with a value close to unity indicating a good estimation. In [3], motion reliability is derived directly from the residual error after motion compensation. More recently, Patras et.al. formulated the confidence measure as a posteriori probability under a Bayesian framework [4].

The method we propose in this paper measures the motion confidence on a video block based on its *phase information* across two frames. More specifically, we show that in a phase-matched error image, a correct estimation induces a significant shift of energy from the lowpass frequencies, and a reverse shift when such block positions do not allow for a reliable estimation. The algorithm is presented in the paper as follows.

Section 2 describes how a phase-match error is constructed. In section 3, we formulate the confidence measure based on this error, and demonstrate its improved performance. Section 4 arrives further at an analytical model for the error and the proposed confidence measure. Finally in section 5 we shows results on some sequences in which areas of unreliable motion information are identified using this measure.

2. THE PHASE-MATCHED ERROR IMAGE

The motion confidence measure is based on a phase-matched error between two frames. As compared with a conventional error which is created by subtracting the matching image from the original, in a phase-matched error the phase component of the matching image is replaced by the phase component of the original image before subtraction. This calculation is depicted in Figure 1 and described below.

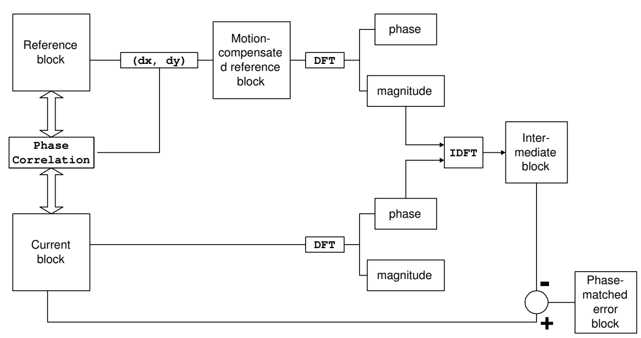


Fig. 1. Calculation of a phase-matched error

Using phase correlation to calculate the interframe displacement [5], a best-matched block is located on the reference frame. The Fourier transform is obtained on this matching block, which consists of the magnitudes and phases. The phases from this transform are then removed and replaced by the phases from transform of the current block. The resulting product is inverse-transformed to produce an intermediate matching block, which is then subtracted from the current block to create the final phase-matched error.

Figure 2 shows an example phase-matched error, calculated between two frames of *Mobile and Calendar*. We aim to evaluate the motion confidence from properties of these errors.

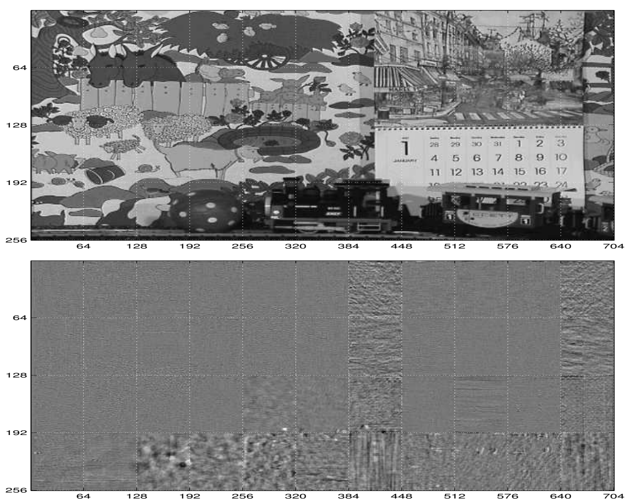


Fig. 2. A frame of *Mobile and Calendar*, and its phase-matched error calculated using block size of 64x64

3. MOTION CONFIDENCE MEASURE

An unreliable estimation could be caused by the presence of multiple motions in the selected area, or a lack of texture for a meaningfull estimation. Nonetheless, we expect a logical decision from a motion confidence, being either

- The motion estimation is reliable, or
- The motion estimation is unreliable

Figure 3 displays cases of good and bad motion confidence. Three blocks are selected on the sequence “Mobile and Calendar”, such as two of them are translating, and the other straddles a motion boundary. Using motion vectors extracted from phase-correlation, Figure 3-b shows the motion compensated blocks, and Figure 3-c shows the corresponding phase-match errors as proposed.

The motion information on the middle block is unreliable, because there are two motions in the scene. Because it is usually necessary to set a threshold on the confidence measure to separate between good and bad estimations, in case of multiple motions it is desirable that such measure would respond to small disturbances in the motion field as well as it would to larger irregularities. In this context, a small disturbance means a second, non-dominant object whose size

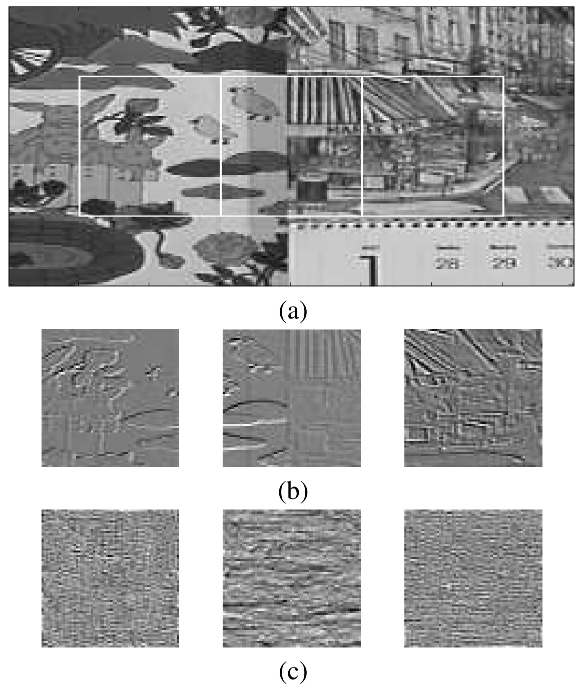


Fig. 3. (a) Three adjacent blocks in “Mobile and Calendar”, (b) Residual errors, and (c) Phase-matched errors

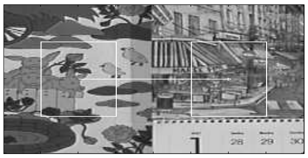
is much smaller than the dominant object, or whose motion is only slightly different from the dominant motion, or a combination of both.

With analytical details to follow in later parts of this paper, we defined our confidence measure $\mathcal{R}$ as:

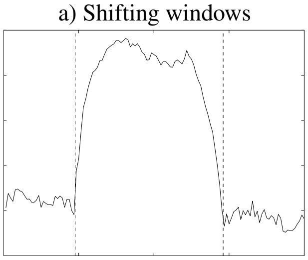
$$\mathcal{R}(E_{pm}) = \frac{\sum (E_{pm}^{lowpass})^2}{\sum E_{pm}^2} \quad (1)$$

in which E_{pm} is the phase-matched error, and $E_{pm}^{lowpass}$ is a low-pass version of this error. The ratio $\mathcal{R}(E_{pm})$ represents the proportion of energy in the lower-frequency components. This measure of reliablity is a function of an input error imageE_{pm}. The lowpass image is calculated using a simple DCT operation as follows. The coefficients of a phase-matched error is calculated first, using the standard DCT transform. We then retain only the lower-frequency components - i.e coefficients occupying the top left triangle, and set the rest of the coefficients to zero. The lowpass image is then produced by an inverse DCT transform.

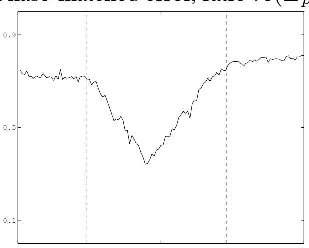
The following experiment is setup to demonstrate the effectiveness of this measure $\mathcal{R}$. In figure 3, imagine that instead of observing three discrete, non-overlapping blocks, we gradually shift the left-most block toward the right, one pixel at a time, until it converges on the right block. Under this transition, the coherency of motion field inside this square window is gradually disrupted as it slides over the boundary between the calendar and the background. The amount of disruption is increasing as the window moves further into the calendar. After the calendar becomes the dominant object in this window, the motion coherency gradually resumes as this window moves in its entirety into the calendar. Under the transition, estimated motions on any block positioned across the boundary are unreliable.



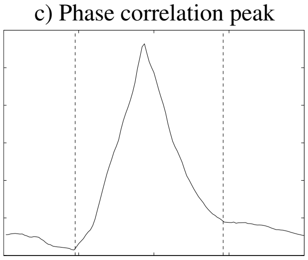
a) Shifting windows



b) Phase-matched error, ratio $\mathcal{R}(E_{pm})$



c) Phase correlation peak



d) Sum-squared error, $\sum E_{mc}^2$

Fig. 4. Comparison with some other confidence measures under a shifting window. The left vertical line marks the starting position from which the motion field has two distinct motions , and the right vertical line marks the position when this disruption ends

The performance of our measure is illustrated in figure 4-b. On this graph, values of the proposed motion confidence measure are plotted at each position as the viewing window moves along. From left to right, the two vertical lines on the graphs mark the positions when the disruption to a single motion field starts and ends. As it is seen in this graph, the confidence measure $\mathcal{R}$ is low in the regions of translated motion, because the motion estimations are correct in these regions. As soon as the viewing windows moves to contain the motion boundary, a sharp rise is observed in the value of $\mathcal{R}$, indicating the sentitivity of this measure when the motion is incorrectly estimated.

This performance is in contrast with two other measures, the sum-squared error in 4-d, and the phase correlation peak in 4-c. In both cases the respective measure changes as the motion estimation becomes unreliable near the boundary region. These changes however appear to depend *linearly* on the amount of disruption, meaning these measures respond poorly to small disruptions and would be undesirable where a binary decision is expected of them.

Besides motion boundaries, blocks with low texture is also an unreliable target for motion estimation. Using $\mathcal{R}$ and the phase-matched error E_{pm}, we define the likelihood of an unreliable motion estimation as:

$$P = \begin{cases} 1, & \text{if } (\mathcal{R}(E_{pm}) > T_1) \text{ or } (\sum E_{pm}^2 < T_2) \\ 0, & \text{otherwise} \end{cases}$$

where T_1 and T_2 are predefined thresholds. T_1 detects block straddling motion boundaries, while T_2 identifies areas of low texture. The value of $P = 1$ corresponds to an unreliable estimation, while $P = 0$ indicates that the estimated motion is a good match.

4. A MODEL FOR THE PHASE-MATCHED ERROR

The previous experiment has yielded that a correct estimation reduces significantly the proportion of lowpass energy, $\mathcal{R}$, in a phase-matched error. In this section we aim to establish the analytical support for this argument.

Let Δ be the residual motion between block i_k and i_{k-1} between two frames. Phase correlation usually obtains a nearest-integer motion estimation, hence this residual can be assumed as being sub-pixel, or $|\Delta| < 1$

$$\begin{aligned} i_{k-1}(x) &= i_k(x+\Delta) \\ &= i_k(x) + \Delta.i'_k(x) + \frac{\Delta^2}{2!}i''_k(x) + \frac{\Delta^3}{3!}i'''_k(x) + h.o.t \\ &\approx i_k(x) + \Delta.i'_k(x) \qquad (2) \end{aligned}$$

Taking Fourier transform of both sides of this equation, we have

$$\begin{aligned} I_{k-1}(w) &= \mathcal{F}(i_k(x) + \Delta.i'_k(x)) \\ &= (1 + \Delta.j.w)I_k(w) \end{aligned}$$

and hence

$$\begin{aligned} |I_{k-1}(w)| &= |I_k(w)|.|1 + \Delta.j.w| \\ &= |I_k(w)|\sqrt{1 + (\Delta.w)^2} \end{aligned}$$

From our definition in section 2, a phase-matched error can be expressed as

$$\begin{aligned} E_{pm} &= i_k - \mathcal{F}^{-1}(|I_{k-1}|.e^{j\theta_k}) \\ &= \mathcal{F}^{-1}((|I_k(w)| - |I_{k-1}(w)|).e^{j\theta_k}) \\ &= \mathcal{F}^{-1}((1 - \sqrt{1 + (\Delta.w)^2}).|I_k(w)|.e^{j\theta_k}) \end{aligned}$$

or

$$\mathcal{F}(E_{pm}) = (1 - \sqrt{1 + (\Delta.w)^2}).|I_k(w)|.e^{j\theta_k}$$
$$= (1 - \sqrt{1 + (\Delta.w)^2}).I_k(w) \quad (3)$$

This equation is of particular interest as it expresses the relationship between the phase-matched error E_{pm} and the transform I_k of the original image block. Function $|(1 - \sqrt{1 + (\Delta.w)^2})|$ is plotted in figure 5 as a function of w and with $\Delta < 1$, together with the linear function $\Delta.w$. From this graph it can be derived that in the phase-matched error E_{pm}, the low-frequency components are significantly suppressed as compared with the original image. It should also be noted that (3) is derived from an approximation of the Taylor's series in (2), under the assumption that the motion has been estimated properly to the nearest integer value. This is evidently demonstrated in figure 4-b, where the proposed measure takes on a low value when the motion is estimated correctly, and a high value otherwise. An inaccurate estimation will cause a mismatch between original and estimated blocks, hence promptly increasing the lowpass frequencies in the error image.

5. RESULTS

Figure 6 shows the performance of this measure on two sequences, *Mobile and Calendar* and *Table Tennis*, with the same threshold settings. In this figure, those areas with low motion confidence is shown as being crossed. In *Mobile and Calendar* the algorithm identifies all block positioned across a motion boundary as having unreliable motion information. The measure is very sensitive, for example blocks on the wallpaper but containing a small part of the locomotive chimney are properly detected. Likewise, on *Table Tennis* all block containg both the background and part of the hand, or the ball, are properly identified.

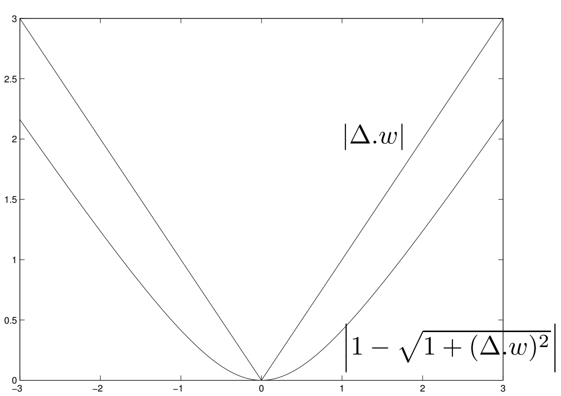


Fig. 5. Transfer function for the phase-matched error

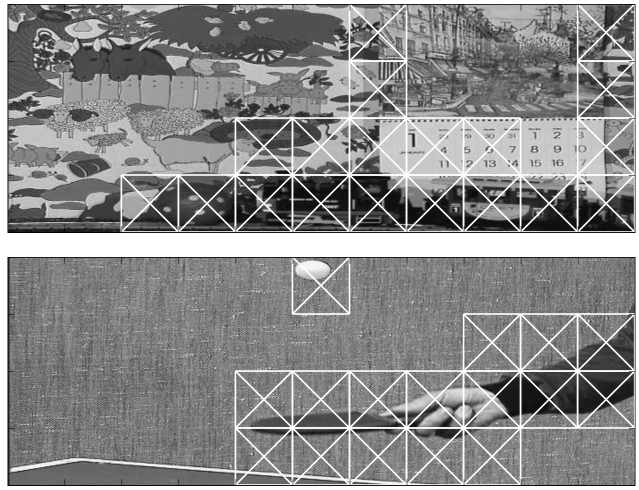


Fig. 6. Areas with unreliable estimations being shown as marked blocks

6. CONCLUSION

In this paper we have introduced a novel motion confidence measure to efficiently identify regions of unreliable motion estimation. In an application such as object segmentation, the results obtained in previous example can be used prior to a spatial segmentation step. Such information on motion reliability also helps an encoder choose optimal block sizes for motion estimation for video coding.

7. REFERENCES

[1] Jianbo Shi and Corlo Tomasi, "Good features to track," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 1994.

[2] Stefan Kruger and Andrew Calway, "Motion estimation and tracking using multiresolution affine models," in *Proceedings of the IEE Colloquium on Motion Analysis and Tracking*, May 1999, pp. 50–55.

[3] Astrid Lundmark, Haibo Li, and Robert Forchheimer, "Motion vector certainty reduces bit rate in backward motion estimation video coding," in *Proceedings of SPIE Visual Communications and Image Processing*, June 2000, vol. 4067, pp. 95–104.

[4] E.A. Hendriks Ioannis Patras and R.L. Lagendijk, "Confidence measures for block matching motion estimation," in *International Conference on Image Processing*. IEEE, 2002, vol. 2, pp. 277–280.

[5] C.D. Kuglin and D.C. Hines, "The phase correlation image alignment method," in *IEEE Conference on Cybernetics and Society*, 1975, pp. 163–165.