

Cepstral Analysis of Optical Flow

by

Esfandiar Bandari
and
James J. Little

Technical Report 92-6

Department of Computer Science
University of British Columbia
Rm 333 - 6356 Agricultural Road
Vancouver, B.C.
CANADA V6T 1Z2

Cepstral Analysis of Optical Flow

Esfandiar Bandari and James J. Little
Technical Report 92-6

November 1991

Laboratory for Computational Vision
Department of Computer Science
University of British Columbia
Vancouver, British Columbia, CANADA V6T 1W5
email: bandari@cs.ubc.ca

Abstract

Visual flow analysis from image sequences can be viewed as detection and retrieval of echoes or repeated patterns in two dimensional signals. In this paper we introduce a new methodology for optical flow analysis based on the *cepstral* filtering method. Cepstral filtering is a non-linear adaptive correlation technique used extensively in phoneme chunking and echo removal in speech understanding and signal processing. Different cepstral methodologies, in particular *power* cepstrum, are reviewed and more efficient variations for real image analysis are discussed.

Power cepstrum is extended to multiframe analysis. A correlative cepstral technique, cepsCorr, is developed; cepsCorr significantly increases the signal to noise ratio, virtually eliminates errors, and provides a predictive or multi-evidence approach to visual motion analysis.

1 Introduction

In biological and computational vision, time-varying image analysis plays a significant role in segmentation of the scene, encoding 3D information, egomotion estimation, object tracking, determination of focus of attention, and estimation of time of collision [Nak85]. As a result, a number of different methodologies for motion analysis have been developed, including gradient based approaches [HS81], correlation matching schemes [Ana89, LBP88, BLP89, Fua91], spatio-temporal analysis [AB85], token matching techniques [MU81] and velocity selective mechanisms [Hee87]. In this paper we introduce a new approach based on the detection and retrieval of echoes from two dimensional signals over time.

A sequence of images over time can be viewed as spatial and temporal echoes. For instance, given two images of a stationary scene the later image can be viewed as a temporal echo of the first. By analogy, visual motion can be viewed as recognition and determination of spatial echo from one frame to the next. Therefore while a stationary scene has a zero interval for its spatial echo, simple translation of the scene (due to lateral motion of the camera) generates a constant spatial echo period across the image plane.

Bogert, Healy and Tukey [BHT62] first introduced *cepstral filtering* and the *que-frency approach* for estimating the arrival time of the echo of a complex signal. The cepstrum of a complex signal was first introduced as the power spectrum of the logarithm of the power spectrum of the signal. Oppenheim [OSS68] developed the complex cepstrum of the signal for detection and removal of echoes, multiplicative filters, predictive deconvolution and inverse filtering of signals. Polydoros *et al* [PAF79] introduced differential cepstrum as a shift invariant complex cepstrum.

Yeshurun and Schwartz [YS89] proposed cepstral filtering for binocular stereo based on the existence of ocular dominance columns in the primary visual cortex; DeYoe and Van Essen [DE88] describe the existence of specialized neurophysiological channels based on the magnocellular cells which carry both stereo and motion information to the middle temporal area. From a computational point of view, there is obviously a great deal of similarity between analysis of motion and stereopsis [Hor86].

In the next section we provide a detailed overview of the power cepstrum technique and a brief review of variations to this approach. In Section 3 we introduce modifica-

tions to the traditional power cepstrum and differential cepstrum schemes to enhance both their computational and their analytical performance for the detection of the echo arrival period. Section 4 presents the results of applying different cepstral filtering techniques to motion sequences. In Section 5 we introduce cepsCorr, a correlative cepstral technique that eliminates possible ambiguities from cepstral analysis and has a high signal to noise ratio [Pra78], in comparison with standard correlation methods.

In summary, we modify power cepstral and differential cepstral methods to increase efficiency and improve the signal to noise ratio. We show how to perform multiframe cepstral analysis. Finally, and most importantly, we develop cepsCorr, which significantly increases the signal to noise ratio, and virtually eliminates errors.

2 Cepstral Filtering: Themes, Variations and Mathematical Preliminaries

In this section we will provide a mathematically descriptive explanation of power cepstral filtering, and a brief review of variations that have been introduced over time, leading to the differential cepstrum analysis. For simplicity the majority of the mathematical derivations and explanatory formulas are provided for one dimensional signals, which are easily generalizable to images or higher dimensional signals.

2.1 Power Cepstrum

Power cepstrum has historically been defined as the *power spectrum of the logarithm of the power spectrum of a signal*. Given a signal $h(x)$ comprised of an original input $s(x)$ and its echo delayed by τ :

$$h(x) = s(x) + s(x - \tau) \quad (1)$$

one can simply construct the Fourier transform of the signal:

$$\begin{aligned} \mathcal{F}\{h(x)\} &= \mathcal{F}\{s(x) + s(x - \tau)\} \\ &= \mathcal{S}(f)(1 + e^{-2\pi i \tau f}) \end{aligned} \quad (2)$$

Taking the logarithm of the magnitude of this complex function:

$$\log(\|\mathcal{H}(f)\|) = \log(\|\mathcal{S}(f)\|) + \log(1 + \cos(2\pi\tau f)) + \text{constant} \quad (3)$$

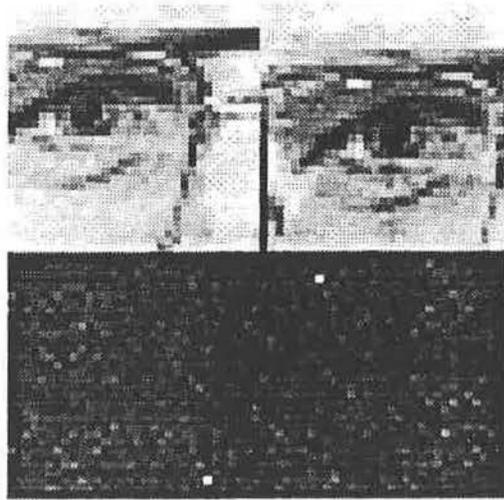


Figure 1: Cepstral analysis of synthetic motion (x,y) displacement is (7,3). The origin of the cepstrum is located at the middle of the top row of the lower portion of the figure.

one observes that $\log(1 + \cos(2\pi\tau f))$ can be expressed as:

$$\sum_{n=1}^{\infty} \frac{(-1)^n \cos^n(\pi f\tau)}{n} \quad (4)$$

The logarithm operation transforms the power spectrum of the signal into the logarithm of the power spectrum of the original input and a residual summation of decreasing cosines generated due to the presence of the delayed replica. Figure 1 shows the power cepstrum of an image and its spatially transformed temporal echo for a simulated motion field of three rows down and seven columns to the right. As expected from the second power spectrum, the cepstrum result is a symmetric function. The location of the two peaks from the center line and the edges of the resulting figure indicate the amount of the horizontal and the vertical motion.

It is important to note that the echo arrival time, τ , is contained in the above summation. To extract this interval, traditionally, a second power spectrum has been applied, which in turn transforms the cosine series into

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \delta(x - n\tau) \quad (5)$$

which is a ripple of decreasing Kroenecker deltas with periodicity τ . The peaks in the cepstrum in Figure 1 reflect the amount of motion between successive frames.

Cepstral analysis behaves as a deconvolving operation – i.e., if two functions are combined by convolution (for example, the input to a filter and the impulse response of a filter), the cepstrum of the resulting signal is equal to the sum of the cepstra of the convolved functions. In fact the echo of a signal is nothing more than convolution of the original signal with a delayed impulse function. Consequently, cepstral filtering can be a powerful technique for dealing with problems involving convolution, deconvolution, and separation of signals. Oppenheim, Shafer and Stockham [OSS68] introduced the complex cepstrum, the inverse z transform of the complex logarithm of the z transform of the signal:

$$\hat{h}(x) = Z^{-1}\{\log(Z\{h(x)\})\}$$

which soon became widely used, especially in speech research, for predictive deconvolution, removal of echoes and inverse filtering.

Given a signal $s(x)$ and its echo $s(x + \tau)$:

$$\begin{aligned} \hat{H}(z) &= \log(H(z)) \\ &= \log(S(z)) + \log(1 + z^{-\tau}) \end{aligned} \tag{6}$$

$$= \hat{S}(z) + \sum_{n=1}^{\infty} \frac{(-1)^n}{n} z^{-n\tau} \tag{7}$$

Taking the inverse z transform, we have:

$$\hat{h}(x) = \hat{s}(x) + \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \delta(x - n\tau)$$

Using the proper comb function with periodicity τ , one can remove the effects of the echo and reconstruct the original signal through inverse filtering.

Complex cepstrum analysis soon found other applications in image processing, such as testing the stability of 2D recursive filters [Dud77].

2.2 Phase Cepstrum

At this point we investigate the mathematical relationship between power and complex cepstrum [SC75] [CSK77]. Evaluating Eq. 7 on the unit circle ($z = e^{2\pi if}$) (i.e., Fourier

transform) one can derive:

$$\begin{aligned}
 \hat{H}(f) &= \log(S(f)) + \log(1 + e^{2\pi i f \tau}) \\
 &= \|S(f)\| + i \text{phase}(S(f)) \\
 &\quad + \frac{1}{2} \log(2 + 2 \cos 2\pi f \tau) \\
 &\quad + i \tan^{-1}\left(-\frac{\sin 2\pi f \tau}{1 + \cos 2\pi f \tau}\right)
 \end{aligned} \tag{8}$$

The first and third term above are part of the familiar power cepstrum technique. Moreover, note that the fourth term of the above equation, which is part of the imaginary portion of the signal, also contains the echo arrival time of the signal, τ . The inverse transform of this imaginary portion, commonly called the *phase cepstrum*, also yields peaks at multiples of the echo arrival time in much the same way that power cepstrum does.

Unfortunately, however, phase cepstrum, like other phase correlation techniques [KH75] is highly susceptible to noise, and like the complex cepstrum can suffer from possible errors caused by phase unwrapping. It is included here primarily for the sake of completeness.

2.3 Differential Cepstrum

If we attempt to use the complex cepstrum the ambiguity of phase introduces an immediate dilemma, namely the output of the system, $\hat{s}(x)$, will no longer be unique. This should be inherently obvious since the logarithm of a complex signal, G , is another complex signal, G' , in which the real part of G' is the logarithm of the amplitude of G , and its imaginary part is the phase of the input signal G . But since the complex cepstrum of an image involves the inverse z transform of the complex logarithm of its z transform, the phase of the complex logarithm should be made continuous and periodic, in both dimensions μ and ν .

One dimensional phase unwrapping algorithms to disambiguate the phase [Tri77] can be extended to two dimensional signals. In fact, Dudgeon derived two methods for calculating the complex cepstrum of two dimensional signals: using the phase unwrapping of complex logarithm, and the recursion relation for minimum phase signals [Dud77]. The phase perplexities, however, greatly reduce the signal to noise ratio and

add additional computational costs. Moreover, almost from the start when complex cepstrum was introduced by Oppenheim as a homomorphic filter for image reconstruction, the derivative of the complex log was used in the analysis.

Noting this, Polydoros, Au and Fam [PAF79] used the logarithmic derivative rather than the plain logarithm of the complex signal for analysis. This technique, named *differential cepstrum*, like its predecessors, still transforms a convolution:

$$h(x) = s_1(x) * s_2(x)$$

into an addition:

$$\hat{H}_d(z) = \frac{\frac{d}{dz}H(z)}{H(z)} = \frac{\frac{d}{dz}S_1(z)}{S_1(z)} + \frac{\frac{d}{dz}S_2(z)}{S_2(z)} \quad (9)$$

But, additionally, since the derivative of the logarithm is the derivative of the signal divided by the signal itself, the differential logarithm is also scale invariant.

Polydoros *et al* also showed that differential cepstrum is also shift invariant; that a delay of the total input sequence does not affect the differential cepstrum except for the first sample, which is indeed a measure of the delay; and that, aside from side-stepping the phase ambiguities, differential cepstrum also provides greater computational efficiency over the complex cepstrum analysis.

Looking at a signal and its echo (Eq. 1), and using the Fourier transform, the logarithmic derivative becomes:

$$\begin{aligned} \hat{H}_d(f) &= \frac{\frac{d}{df}S(f)}{S(f)} - \frac{2\pi i\tau e^{-2\pi i\tau f}}{1 + e^{-2\pi i\tau f}} \\ &= \hat{S}_d(f) - \pi i\tau - \pi\tau \tan(\pi\tau f) \end{aligned} \quad (10)$$

The echo arrival time, τ , is again encoded as a parameter to the tan function and it can be retrieved using the inverse Fourier transform, producing a rippling effect with spatial periodicity τ .

Raghuramireddy and Unbehauen [RU85] provided a simple extension of the one dimensional differential cepstrum to images. This is the sum of the logarithmic derivatives along both axes:

$$\hat{H}_d(z_1, z_2) = \frac{\frac{\partial}{\partial z_1}H(z_1, z_2)}{H(z_1, z_2)} + \frac{\frac{\partial}{\partial z_2}H(z_1, z_2)}{H(z_1, z_2)} \quad (11)$$

It is important to note however, that if Fourier transform is used, this definition eliminates the conjugate symmetry which results from the transformation of real images. Consequently, the inverse Fourier transform of this logarithmic derivative of the Fourier transform of an image, unlike the power or complex cepstrum, does not result in a real signal. Moreover, if the direction of motion is known *a priori*, then the directional derivation can improve the signal to noise ratio of the resulting signal.

3 Improvements to Classical Formulation: Cepscos and DiffcepsSin

In Eqs. 3 and 4, the presence of echo in a signal manifests itself as $\log(1 + \cos(2\pi\tau f))$ which is a *even real function*; the additional $\log(\|\mathcal{S}(f)\|)$, therefore, reduces the signal to noise ratio of the complex cepstrum. Lee, Mitra and Krile [LMK89] provide modifications to the power cepstrum that eliminates the cepstrum of the echoless signal.

Moreover, since we are primarily interested in the effects of $\log(1 + \cos(2\pi\tau f))$, it seems natural to use the real part of the Fourier transform of the logarithm of the power spectrum of a signal and its echo, instead of the power spectrum of the logarithm of its power spectrum. This modification eliminates the effects of the odd parts of the $\log(\|\mathcal{S}(f)\|)$, and reduces the computational overhead of calculating the power spectrum from the Fourier transform. Astute readers will note that even though this modification enhances the signal to noise ratio of complex signals and their echoes, for an image consisting of real values, the power spectrum, and therefore its logarithm, is also an even real function. Thus, even though the new power cepstrum theoretically should not produce different results for real signals, it does increase the computational efficiency and reduce the noise caused by numerical errors.

The reduction in computation time of Cepscos is two fold. Firstly, the second Fourier transform is replaced by the calculation of the discrete cosine transform [Rao90]. Secondly, the calculation of the power spectrum from the Fourier transform (i.e., two multiplications, an addition and a square root) is replaced by an absolute value.

Using the same methodology, one finds even greater improvements for the differential cepstrum technique. A closer look at Eq. 10 immediately reveals that greater efficiency and accuracy could be achieved through the use of imaginary portion of the

inverse Fourier transform of the real part of the logarithmic derivative of the Fourier transform of the signal and its echo. This is equivalent to calculating the Sine transform of the real portion of the logarithmic derivative.

To distinguish more clearly between the the traditional power and differential cepstrum, and our modified methodologies, we named our techniques CepsCos and DiffCepsSin respectively.

4 Results and Comparative Analysis With Standard Techniques

In this section, we present the results of application of cepstral filtering to both synthetic and genuine motion fields. Our main objective is to provide practical examples of the applications, as well as the limitations, of these schemes. Furthermore, the relationship between cepstral filtering and other more traditional approaches to optical flow analysis is discussed, and its strengths and weaknesses with respect to these techniques is examined.

4.1 Applications

The following examples provide the results of applying differential and power cepstrum techniques to synthetic and actual motion fields. Figure 1 provides an example of the application of power cepstrum to a synthetic motion field of $dx = 7$ and $dy = 3$. Similar to Yeshurun and Schwartz's approach [YS89], the two frames are placed adjacent to one another, rather than generated by simple addition. Unlike their work, however, our choice was not motivated by the neurophysiological construction of the ocular dominance columns, but by practical signal processing considerations. Although this format of processing increases the computational cost of the Fourier transforms involved, it also increases the resulting signal to noise ratio by moving the cepstral peak away from the corners of the image, and thus reducing the deterioration of the signal by aliasing frequencies.

Consequently the image and its temporal echo will have the form:

$$h(x, y) = s(x, y) + s(x - w - dx, y - dy)$$

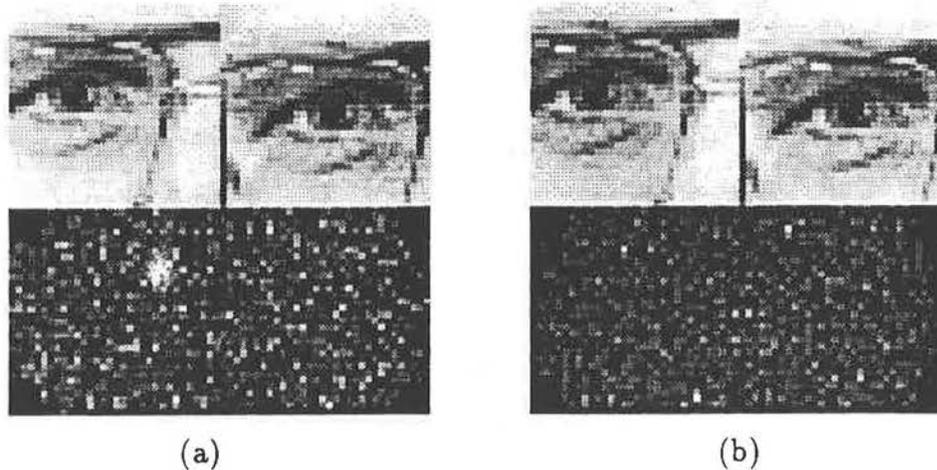


Figure 2: (a) Differential cepstrum of synthetic motion. (b) Differential sine cepstral analysis of synthetic motion. Note the improved signal to noise ratio in (b)

and its cepstrum:

$$\hat{h}(x, y) = \hat{s}(x, y) + \sum_{n=1}^{\infty} \delta(w \pm ndx, \pm ndy)$$

where w is the width of each spatial window. The peak signal will occur in the cepstral plane, between columns $w/2$ and $3w/2$. Figure 1 shows an example of the cepstral peak and its accurate detectability, with dx being measured from the center, and dy from the horizontal edges.¹

Figure 2 shows the same signal analysis utilizing the differential cepstrum and differential sine cepstrum filters.² The differential sine cepstrum technique is much more effective in motion detection than the traditional differential cepstrum, and computationally less expensive than the power cepstrum. However, our experience has shown that the computational efficiency achieved by differential sine cepstrum does not balance the reduction in the signal to noise ratio relative to power cepstrum.

Figure 3 shows the result of the power cepstrum on a natural outdoor scene; the

¹The corner regions of the cepstral image are inconsequential to the cepstral peak detection, and were therefore reduced to zero to increase the scaling of the cepstral peak in the image.

²In these figures the absolute value of the signal rather than the actual signal is displayed. Unlike the power cepstrum, for the differential cepstrum, the location of the signal is not detected merely by its spatial location, but also by the sign of the peak.

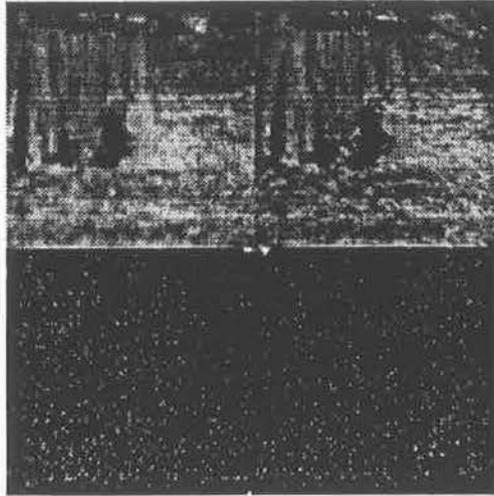


Figure 3: Cepstral analysis of a natural scene

two symmetric peaks are located at the correct motion. Figure 4 shows the optical flow field of selected points due to self motion, calculated using power cepstral filtering. Note that, as expected, the motion of distant objects is less than the apparent motion of objects close to the moving observer.

The gray arrows show where the DC component of the cepstrum had to be taken into account. These are generally in the areas of low contrast or areas of close approximation to the occluding motion boundaries; more will be said on this in 4.3.

4.2 Multiframe analysis

Modeling motion as spatial and temporal echoes naturally suggests extension of cepstral techniques to multiple frame analysis. A simple mathematical derivation for constant velocity, or accelerating motion, shows that the resulting signals are simply the sum of individual echoes between every pair of frames. In the case of constant velocity, the magnitude of echo between successive frames indeed diminishes. This is primarily due to cancellation from positive and negative terms in the Fourier series expansion of the resulting geometric series. However, the echoes have a systematic arrangement across multiple frames and that regularity can be used to improve detection of constant velocity over several frames.



Figure 4: The optical flow field resulting from egomotion

Figure 5 shows the multi-ceps analysis of a simulated motion: the sequence of images creates an input to the cepstral filter. As the (symmetric) echoes indicate, the third frame has a motion of three pixels in the horizontal and vertical direction with respect to the first frame. And the second figure has moved one pixel horizontally and one pixel vertically with respect to the first. The last echo is the result of the motion of the last frame with respect to the second frame (two pixels horizontally and two pixels vertically), displaying an acceleration in the direction of motion.

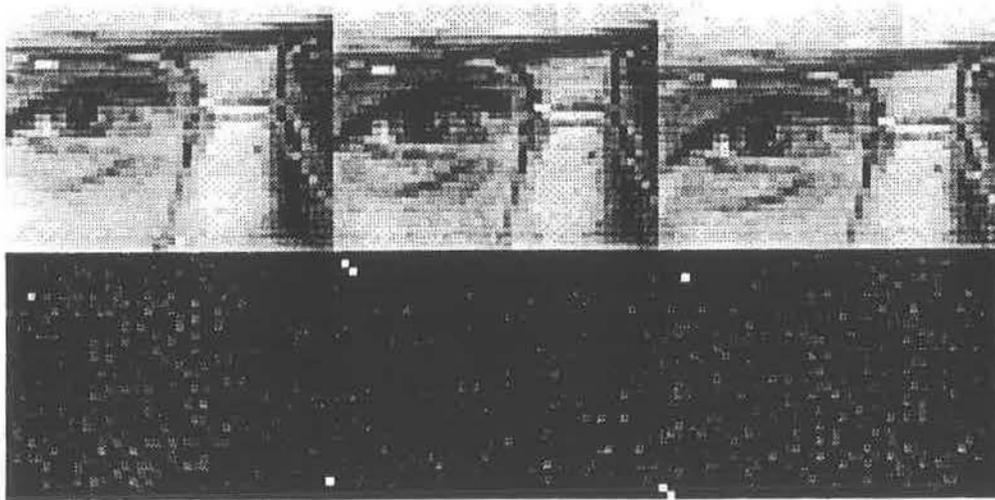


Figure 5: Multi-frame optical flow field analysis utilizing cepstrum

4.3 Analysis of Results and Comparative Study with Standard Techniques

At this point it is necessary to discuss the limitations of cepstrum techniques and compare them to other standard computational approaches as well as psychophysical results of human motion analysis.

An obvious question to raise is the size of the cepstral window. Similar to correlation analysis, this size depends on the image content and the magnitude of the motion field. The size of the window is the size of some appropriate image subsection (depending on image content) plus margins that are equal in width to the magnitude of the expected maximum motion. By now it should be apparent that the cepstral peak occurs in the middle section of the result. Moreover, at margins of the concatenated inputs, new or missing data appear in the figure. The size of these margins influences the signal to noise ratio of cepstrum. We will shortly propose a solution that limits the effects of margins (Section 5).

A less obvious and more challenging characteristic of the cepstral technique is the effect of the DC component, or the overshadowing of the motion peak by the peak at location (0,0) (see Figure 6). The peak at (0,0) is partially caused by the *constant* term in Eq. 3, as well as the contributions of the original signal. Since the location of

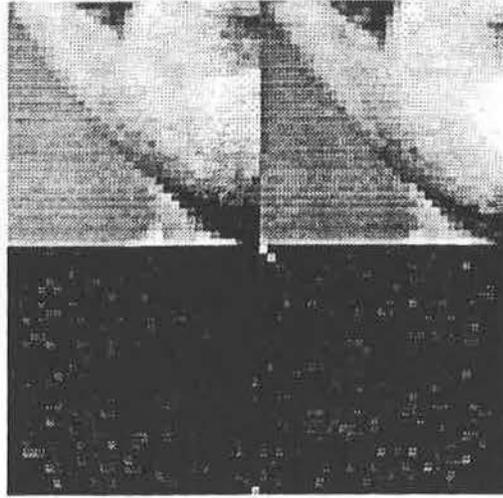


Figure 6: Peak modulation by the DC component. The second highest peak at (1,1) corresponds to the correct motion.

the false peak is predetermined, this problem could be trivially addressed by simple neighborhood checking and thresholding of the ratio of the secondary to the peak magnitude at (0,0). We introduce below a modified technique, *cepsCorr*, to take advantage of the strength of the DC component in motion analysis.

To further study this phenomenon, Figure 7 shows the effects of the false peak in a simulated motion field. The lower left square shows that the cepstrum technique accurately predicts the motion field for the majority of the pixels in the outlined square (the white pixels). The gray pixel represents the location of the only pixel that was affected by the false peak and whose calculated optical flow was subsequently corrected. The black pixel represents the motion field that was calculated erroneously independent of the DC effect. On the other hand, the gray pixel in the lower right square shows that even though this predicted motion was erroneous, it was within one pixel of the actual simulated motion.

Moreover, the adverse effects of the high frequency portion of the original signal can be easily reduced through simple subtraction of the cepstrum of the original signal [LMK89]. We have also experimented with multiplying the merged windows with a windowing filter (of the form $1 + \cos(x)$), to reduce the windowing effect, and have

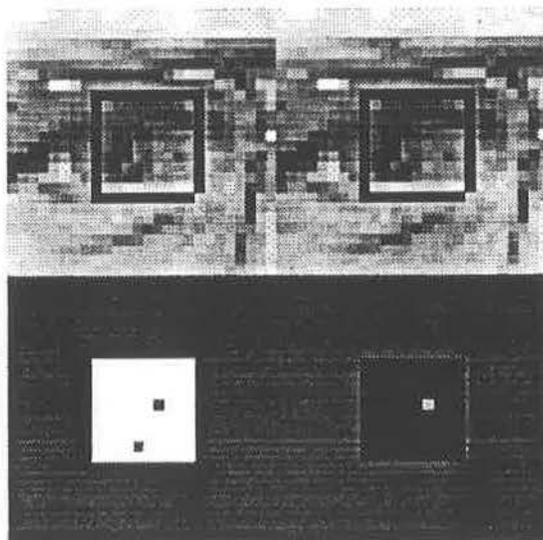


Figure 7: Effects of DC component on correct optical flow estimation. The squares in the top images outline the centers of cepstral filtering windows. Lower left: white area shows the locations of correct estimates and gray identifies the points which are correctable after removal of the DC component, black: incorrect; Lower right: the gray area shows the incorrect estimates that are off by only one pixel.

achieved a reduction in the DC component.

When the portion of the image being analyzed has multiple motions, the output of cepstral filtering will also depict multiple peaks. Figure 8 shows a person moving by between one and two pixels vertically, in front of a static background. The cepstral output displays two bright spots at top and bottom of the lower region for the foreground motion and a strong peak at (0,0) for the static background motion.

From a computational standpoint, as noted by Olson and Coombs [OC90], cepstral filtering can be viewed as a non-linear adaptive auto-correlation technique, with the logarithm function acting as part of the adaptive prefilter to the correlation.

$$\log \|\mathcal{S}(f)\|^2 = \left\| \frac{\sqrt{\log \|\mathcal{S}(f)\|^2}}{\|\mathcal{S}(f)\|} \right\|^2 \mathcal{S}(f) \mathcal{S}^*(f) \quad (12)$$

This prefiltering has a compressive equalizing effect in the frequency domain, reducing the contributions of the narrowband signals and therefore emphasizing the effects of

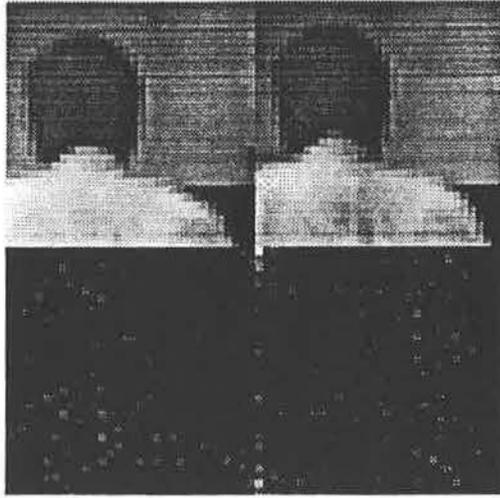


Figure 8: Result of multiple motion in cepstrum window. Both background and figure show signals.

broadband components of the image [OC90]. It is the narrowband signals (i.e., uniform or regular periodic patterns) that increase matching ambiguity and hence reduce the signal to noise ratio in autocorrelation, making the cepstrum a more powerful discriminant and matcher.

This compressive effect in the Fourier domain can be exploited only if the phase of the signal is preserved, leading to the comparison of the cepstrum approach to phase correlation [KH75]:

$$e^{-2\pi i \tau f} = \frac{S_t(f)S_{t+\delta t}^*(f)}{\|S_t(f)S_{t+\delta t}^*(f)\|} \quad (13)$$

Even though the phase correlation technique provides a simple approach to optical flow estimation, it is much more susceptible to random or quantization noise than power cepstrum [LMK89].

5 CepsCorr: A Correlative Use of the Cepstrum

As Figure 4 indicates, cepstral analysis of real image sequences provides accurate estimates of the actual motion. It is however both interesting and desirable to overcome

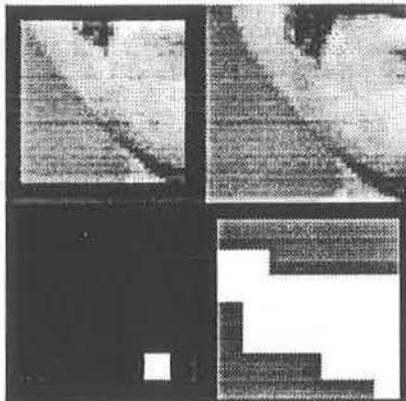


Figure 9: *cepsCorr*: lower right shows in white areas where $(0,0)$ peak is maximal, lower left shows magnitudes of those peaks. Largest peak identifies correct motion.

some of the challenges posed by cepstral filtering. In particular the presence of the DC component proves to be an important factor, one that has to be addressed if we want to distinguish between truly stationary portions of an image and moving ones that are overshadowed by this effect.

A simple approach is to threshold the image based on the ratio of the first and second peak when the maximum peak is at $(0, 0)$. As we mentioned earlier, subtracting the cepstrum of one image and using windowing filters have both increased the signal to noise ratio and reduced the DC component. But thresholding techniques are rather arbitrary and do not address questions such as analyzing the motion of periodic patterns.

A more astute approach is to make the DC component aid us in the correct detection of the motion in the few ambiguous situations, as well as utilizing the cepstrum to provide a degree of confidence or certainty about the inferred measurements. To this end we proceeded with *cepsCorr*, a correlative approach to cepstrum, where the window from one image is used in conjunction with a region of the second to measure motions between the two frames. Essentially, cepstral analysis replaces the sum of squared differences step often used in correlation motion or stereo. The motion chosen represents the displacement for which the cepstral peak at $(0,0)$ is maximum. Figure 9 shows the results of this approach, again on a simulated motion.

The white area in the lower right corner shows locations in the right image where



Figure 10: Performance of *cepsCorr* in Gaussian noise. The black points in lower left figure show the location of the points where the *cepsCorr* estimate is incorrect. The white points in the lower left figure show the location of the points where the motion estimate is off by one pixel.

the DC component of the cepstrum with the left image was in fact the dominant peak. The lower left corner shows the magnitude of these peaks; the bright spot is at the correct displacement. This peak provides a strong discriminant for the motion. In *cepsCorr*, this analysis is performed at each image point, for suitable displacements.

Even on a small correlative window over an ambiguous region, the signal to noise ratio was roughly 25. In our experiments, the majority of the signal to noise ratios were above two orders of magnitude. Obviously, complete correlation is not necessary if the initial peaks of the first few iterations agree on the correct motion field. Several “voting” schemes, based either directly on the magnitude of the cepstral output, or their

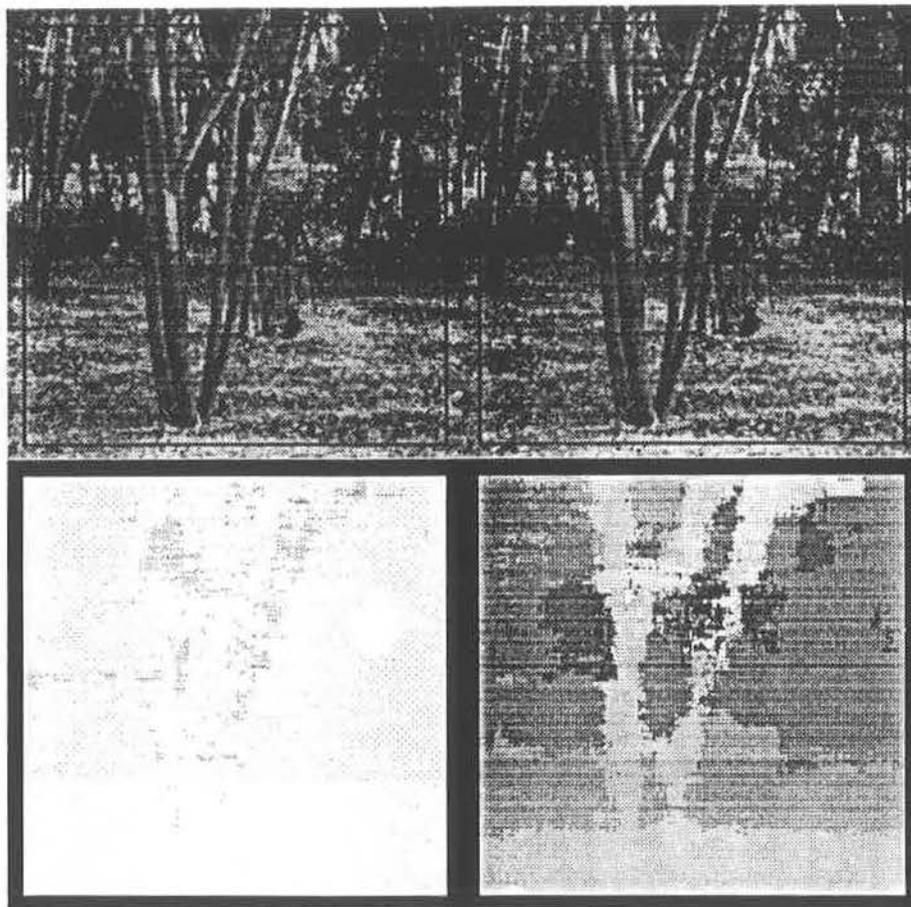


Figure 11: Horizontal and vertical disparity fields for a natural scene using cepsCorr

signal-to-noise ratios, or a combination, immediately suggest themselves. Presently we are using the maximum (0,0) amplitude as the output of cepsCorr.

To test the limits of our work we increased the displacement and introduced Gaussian distributed noise with variance 25. As Figure 10 indicates, cepsCorr performs very strongly in the presence of noise, with the exception of a few points where the image lacks matching features.

CepsCorr was used to analyze the natural scene in Figure 4. From knowledge of the direction of egomotion, the motion search region was restricted to the range of ± 1 in the vertical, and -2 to 7 in the horizontal direction. The output of cepsCorr is shown

in Figure 11. The lower left corner image shows the measured vertical displacement and the lower right displays horizontal displacement. Lighter shades depict larger displacement, which correspond to closer distances. The points in distant horizon that do not go through any significant motion are clearly marked as darker regions.

We also used `cepsCorr` with the two images supplied in the reverse order (of course changing limits of the correlation window to reflect this fact). Unlike regular correlation or sum of the square differences technique, the two results for `cepsCorr` were basically identical, reflecting the robustness of the general technique.

Moreover a preliminary analysis shows that `cepsCorr` provides promising results at motion boundaries. In conjunction with the reduction in signal to noise ratio, this effect should provide a strong clue toward the location of occluding edge contours.

6 Concluding Remarks

In this paper we have introduced a novel nonlinear technique for optical flow analysis based on cepstral filtering, i.e., identifying motion (or stereo) displacements by detection of spatial echoes between successive frames. We introduce modifications to the traditional power cepstral and differential cepstral methods to increase efficiency and improve the signal to noise ratio. We further extended this work to perform multiframe cepstral analysis. Finally, and most importantly, we develop `cepsCorr` – a correlative approach using cepstral analysis – which significantly increases the signal to noise ratio and provides a predictive or multi-evidence approach to visual motion analysis.

There are many other interesting questions that can be studied, including: multi-scale analysis, motion estimation to subpixel accuracy, integration of various cepstral techniques, localization of occluding contours, improvements of signal to noise ratio, the effects of prefiltering and various noise, and detection of contour symmetries. We have implemented segmentation of stationary textures based on cepstral filtering and would like to investigate further the effects of, and the solutions to, the affine transformation of these textures.

One obvious but noteworthy attribute of cepstral filtering is its extensibility to parallel implementation. We have implemented one of our programs on a network of Transputers and would like to address that issue more fully, in addition to reductions in computational costs, and standard techniques to improve performance, such as utilizing

of the Hartley transform.

7 Acknowledgement

We would like to extend our gratitude to David Lowe who initially encouraged this work and whose interest and helpful comments have been invaluable.

References

- [AB85] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, 2(2):284–299, 1985.
- [Ana89] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2:283–310, 1989.
- [BHT62] B.P. Bogert, M.J.R. Healy, and J.W. Tukey. The queferency analysis of time series for echoes: Cepstrum, psuedo-autocovariance, cross-cepstrum and saphe cracking. *Proc. of Symp. On Time Series Analysis*, 1962.
- [BLP89] H. Bulthoff, J. J. Little, and T. Poggio. A parallel algorithm for real-time computation of optical flow. *Nature*, 337:549–553, February 1989.
- [CSK77] D.G. Childers, D.P. Skinner, and R.C. Kemerait. The cepstrum: a guide to processing. *Proc. IEEE*, 65(10), October 1977.
- [DE88] E.A. DeYoe and D.C. Van Essen. Concurrent processing streams in monkey visual cortex. *TINS*, 11(5), 1988.
- [Dud77] D.E. Dudgeon. The computation of two-dimensional cepstra. *IEEE Trans. Acoust., Speech, and Signal Processing*, ASSP-25(6):476–484, December 1977.
- [Fua91] Pascal Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. Technical Report 1369, INRIA, January 1991.

- [Hee87] David J. Heeger. Optical flow from spatiotemporal filters. In *Proc. 1st International Conference on Computer Vision*, pages 181–190, London, UK, 1987.
- [Hor86] Berthold K. P. Horn. *Robot Vision*. MIT Press, Cambridge, MA, 1986.
- [HS81] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [KH75] C.D. Kuglin and D.C. Hines. The phase correlation image alignment method. *Proc. IEEE Int. Conf. on Cybernetics and Society*, pages 163–165, 1975.
- [LBP88] James J. Little, Heinrich H. Bülthoff, and Tomaso Poggio. Parallel optical flow using local voting. In *Proceedings of the International Conference on Computer Vision*, pages 454–459, Tarpon Springs, Florida, December 1988. IEEE, Washington, DC.
- [LMK89] D.J. Lee, S. Mitra, and T.F. Krile. Analysis of sequential complex images, using feature extraction and two-dimensional cepstrum techniques. *J. Opt. Soc. Am.*, 6(6), 1989.
- [MU81] D. Marr and S. Ullman. Directional selectivity and its use in early visual processing. *Proc. R. Soc. Lond. B.*, 211:151–180, 1981.
- [Nak85] K. Nakayama. Biological motion processing: a review. *Vision Research*, 25:625–660, 1985.
- [OC90] T.J. Olson and D.J. Coombs. Real-time vergence control of binocular robots. Technical Report TR 348, University of Rochester, June 1990.
- [OSS68] A.V. Oppenheim, R.W. Shafer, and T.G. Stockham. Non-linear filtering of multiplied and convolved signals. *Proc. IEEE*, 56:2136–2191, August 1968.
- [PAF79] A. Polydoros, W. Au, and A. Fam. Shift invariant homomorphic filtering. In *Proc. of twenty second mid-west symposium on circuits and systems*, 1979.
- [Pra78] W. K. Pratt. *Digital Image Processing*. John Wiley & Sons, New York, NY, 1978.

- [Rao90] K. R. Rao. *Discrete cosine transform : algorithms, advantages, and applications*. Academic Press, Boston, Mass, 1990.
- [RU85] D. Raghuramireddy and R. Unbehauen. The two dimensional differential cepstrum. *IEEE Trans. Acoust., Speech, and Signal Processing*, ASSP-33(4):1335–1337, October 1985.
- [SC75] D.P. Skinner and D.G. Childers. The power, complex and phase cepstra. In *Proc. National Conf. on Telecommunication*, December 1975.
- [Tri77] J.M. Tribolet. A new phase unwrapping algorithm. *IEEE Trans. Acoust., Speech, and Signal Processing*, ASSP-25(1):170–179, December 1977.
- [YS89] Y. Yeshurun and E.L. Schwartz. Cepstral filtering on a columnar image architecture: a fast algorithm for binocular stereo segmentation. *IEEE PAMI*, 11(7), July 1989.