Multi-evidential Correlation & Visual Echo Analysis

by

Esfandiar Bandari
and
James J. Little

Technical Report 93-1

Laboratory for Computational Vision
Department of Computer Science
University of British Columbia
6365 Agricultural Road
Vancouver, British Columbia
CANADA V6T 1Z2

Multi-evidential Correlation & Visual Echo Analysis

Esfandiar Bandari and James J. Little
Technical Report 93-1

Januaray 1993

Laboratory for Computational Vision
Department of Computer Science
University of British Columbia
Vancouver, British Columbia, CANADA V6T 1Z2
email: bandari@cs.ubc.ca

Abstract

Visual motion, stereo, texture, and symmetric boundaries are all repetitions of similar patterns in time or space. If one views these repetitions as "echoes", then the measurement of disparities, segmentation of textons, or detection of boundary symmetries translates into detection of echo arrival periods.

Cepstral filtering, as well as Polyspectral techniques and waveform analysis, are some of the techniques used successfully for echo detection. This paper examines the application of cepstral analysis to computational vision, introduces improvements to the traditional methods, and provides a comparison with other routines presently used.

Finally, we introduce a general multi-evidential correlation approach which lends itself to several computational routines. *CepsCorr*, as we call it, is a simple general technique that can accept different matching routines, such as cepstrum and/or phase correlation as its measurement kernel. The evidence provided by each iteration of cepsCorr can then be combined to provide a more accurate estimate of motion or binocular disparity.

1 Introduction

Webster defines echo as: "1a: the repetition of a sound caused by reflection of sound waves, ... 2a: a repetition or imitation of another: REFLECTION". One can apply a similar abstraction to several common visual routines. For instance, video frames of a stationary scene (i.e., a scene void of moving objects/observer) can be viewed as identical echoes of one another in time. As for moving objects, each object in turn generates an echo or replication of itself on the spatial axes. Consequently, visual motion analysis can be conveyed as detection and determination of echo arrival periods in time and space. This viewpoint is different but analogous to Adelson and Bergen's [AB85] introduction, and Heeger's [Hee87] subsequent use of the space-time cube for spatio-temporal filtering of visual motion.

This conceptual framework can be further extended to other visual tasks such as binocular and trinocular stereo measurements, texture analysis, and detection of boundary symmetries.

Adopting this perspective provides a wealth of background research from other scientific fields - such as acoustics, communication, geophysics, medicine, oceanography, speech understanding, seismology, power engineering and signal processing - which extensively utilize echo detection and removal. In fact, many studies have been conducted in echo analysis and to date the most effective and powerful tool developed has been Cepstrum filtering. [1]

Lee, Krile and Mitra [LKM87] were one of the first researchers that used cepstrum for computational vision and specifically motion analysis. Inspired by the columnar structure of the primates' ocular dominant columns, Yeshurun and Schwartz[YS89] postulated a plausible theory for the neurophysiological mechanism of stereo vision. Olson and Coombs[OC90] then used cepstrum for stereo convergence. Earlier improvements to computational and detection performance of cepstral techniques and their extension to additional areas of computational vision was first reflected in [BL92].

In this paper we examine the application of cepstrum to a few domains in computational vision, particularly emphasizing visual motion analysis. The next section introduces the mathematical and historical background behind cepstrum analysis. Section 3 introduces improvements to the classical power cepstrum which we deem important,

[1] For a sample of the work done using cepstral techniques see [CSK77].

and section 4 applies the cepstrum to texture, trinocular stereo, and multi-frame analysis for motion. Section 5 looks at the symmetric boundary analysis, and section 6 compares the similarities and differences of cepstrum and other linear and non-linear methods in computational vision and digital signal processing.

Section 7 looks at the limitations of the cepstrum approach and tries to address some of the commonly asked questions about its behavior.

Section 8 introduces a simple new approach to detection and matching. CepsCorr is a multi-evidential detection and analysis technique that can be used in conjunction with relaxation labeling, Bayesian analysis, and evidential reasoning. CepsCorr should be viewed as a general framework with cepstrum as a replaceable kernel for detection of single evidences.

2 Mathematical Preliminaries:

Mathematics not only defines and explains the properties of a technique formally, it also furnishes the clues and specifics to improving its behavior. In this section we provide a mathematical description of *power* cepstrum. A brief chronological description of variations to power cepstrum is also provided.

For clarity, the mathematical derivations are written for one dimensional signals, and the motion sequences used in the examples below are generated synthetically.

2.1 Power Cepstrum:

Bogert, Healy, and Tuckey[BHT62] introduced cepstral filtering and *queferency* analysis for estimating the arrival time of the echo of a complex signal. They defined (power) cepstrum as the *power spectrum of the logarithm of the power spectrum* of the signal. Given a signal $h(x)$ consisting of an input $s(x)$ and its echo delayed by τ:

$$h(x) = s(x) + s(x - \tau) \tag{1}$$

and its Fourier transform,

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

the logarithm of its power spectrum is:

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

The Fourier (or more appropriately the Cosine) series of $\log(1 + \cos(2\pi\tau f))$ yields:

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

In effect, the logarithm operation transforms the power spectrum of $h(x)$ into the logarithm of the power spectrum of the single input, $s(x)$, and a residual summation of decreasing cosine polynomials generated due to the presence of the delayed replica.

In addition, the logarithm isolates the echo arrival period, τ, to the above summation of cosines only. To extract this interval, traditionally, a second power spectrum has been applied, which in turn transforms the cosine series into

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

i.e., a ripple of decreasing Kroenecker deltas with periodicity τ.

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.[2]

Other researchers, notably Oppenheim[OSS68] and Polydoros et. al. [PAF79], introduced variations and extensions to power cepstrum that soon found applications in other areas such as non-linear deconvolution, signal decomposition, adaptive filtering, and development of stable IIR filters. Bandari and Little [BL93, BL92] discuss the mathematics of cepstrum variations and their application to computational vision, and introduce computational and performance improvements for some of these techniques.

In our work we have examined different cepstral routines and found power cepstrum to be the simplest and least noisy. But as we will see in section 7, power cepstrum suffers from ambiguities that could be overcome by other cepstral routines.

In the next section we present *new* improvements to power cepstrum which we hope will have wide ramifications for other scientific areas utilizing it. We will also discuss the limitations and advantages of each technique where appropriate.

[2]In these examples, for the sake of clarity, we concatenated the images instead of simply adding them. Therefore the horizontal disparity is increased by the width of the window under consideration. Another beneficial effect of concatenating the images is that it moves the peak of the cepstrum away from the corners of the cepstrum plane, where the noise due to aliasing is largest.

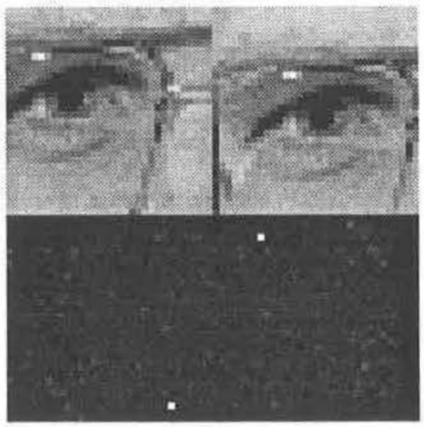


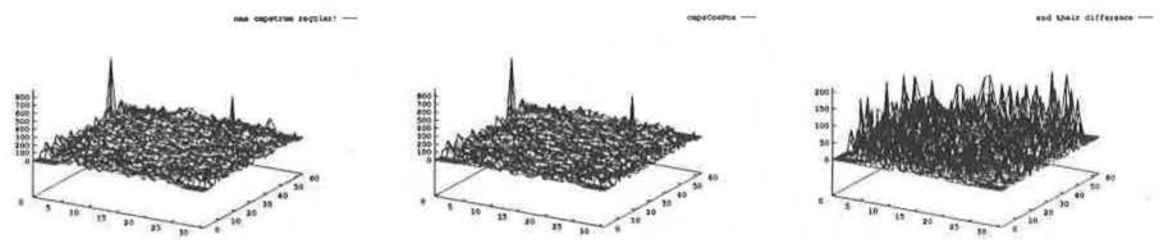
Figure 1: Cepstral analysis of synthetic motion with (x,y) displacement of (7,3). The lower portion of the figure displays the cepstral result with its origin located at the middle of the top row. As expected from the second power spectrum, power cepstrum is a radially symmetric function.

3 Positive CepsCos:

Since its inception, cepstral techniques have become a central tool in many scientific areas such as phoneme chunking, radar and sonar processing, seismology, and medicine. Because of its wide applicability even recently researchers have been pursuing improvements to cepstrum[WY91].

In this section we present changes to power cepstrum that not only improve its signal to noise ratio (SNR), but also increase its computational efficiency.

Reviewing equation 3 it is clear that the echo arrival period, τ, is only present in the term $\log(1+\cos(2\pi\tau f))$. Hence, an obvious but costly method of improving cepstrum, whenever possible, is to subtract the cepstrum of the original signal $s(t)$ from the result [YS89] [LMK89]. In many applications (such as phoneme chunking), however, the form of $s(t)$ is unknown rendering this approach inviable. Moreover, our experience with cepsCorr has shown that subtracting the cepstrum of $s(x)$ can introduce additional noise and ambiguities near the peak value, primarily due to ringing and windowing effects.



(a) (b) (c)

Figure 2: a) regular power cepstrum for Figure 1; b) its positive CepsCos – i.e., cepsCos with all the negative values discarded c) the reduced noise for positive cepsCos

But, unlike $\log(\|\mathcal{S}(f)\|)$, $\log(1 + \cos(2\pi\tau f))$ is *always* a periodic *even* function. Moreover, from equation 5 one observes that the main peak of the power cepstrum is a *positive* δ. Therefore to eliminate the additional noise introduced by the odd portion of the function $\log(\|\mathcal{S}(f)\|)$, as well as all the negative components of its even portion, we introduce *cepsCos*. In short, cepsCos replaces the second power spectrum of cepstrum by its cosine transform and optionally discards the negative portion of the result – i.e., positive cepsCos. By making this simple change, cepsCos reduces the computational cost and increases the signal to noise ratio of the final result.[3]

Figure 2 shows the level of noise reduction between cepstrum and the positive cepsCos for figure 1. As a quantitative measure the signal to noise ratio[4] increased from 20.961 for regular power cepstrum to 36.132 for positive cepsCos and to 130.829 for regular cepsCos. A more important reason for using positive cepsCos, as we will see, is to eliminate the possible false signals affecting subpixel measurements.

DiffCepsSin uses a similar analysis to improve the behavior of differential cepstrum. For the interested reader, the mathematics of this improvement is discussed in [BL93].

In the remainder of this work cepstrum refers to cepsCos unless otherwise noted.

[3] In image processing the original form of the signal $s(x)$ is real – i.e., non complex. Consequently, its power spectrum and the $\log(\|\mathcal{S}(f)\|)$ are also even functions. So if for any reason the traditional cepstrum result for a real function is more desirable, one only has to put the result of cepsCos through an absolute value filter and still benefit from the lower computational cost.

[4] Defined as the ratio of average signal to average noise

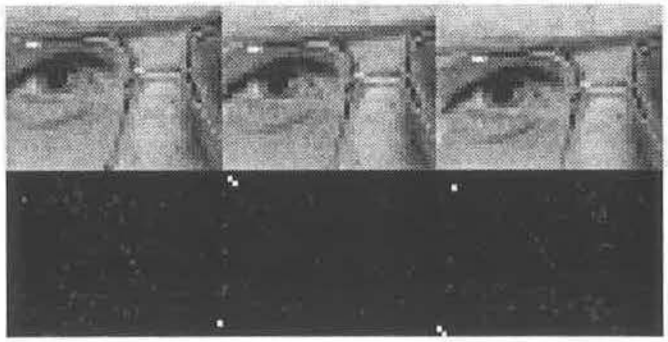
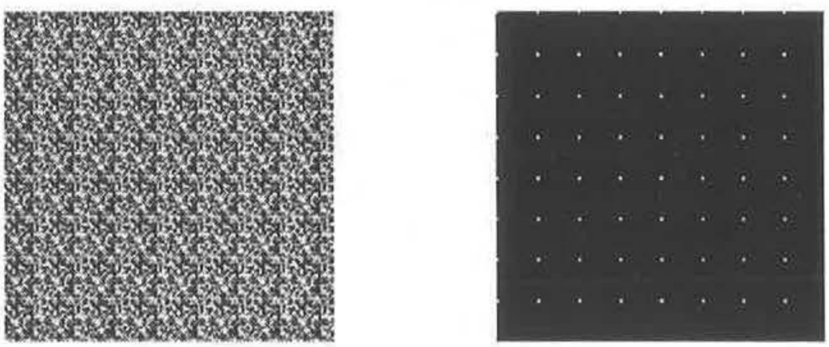


Figure 3: Multi-frame optical flow field analysis utilizing cepstrum. The simulated motion field is made to accelerate; the velocity between frames 1 and 2 is $(1,1)$ while between frames 2 and 3 it is increased to $(2,2)$



(a) (b)

Figure 4: Stationary texture segmentation using cepstrum

4 MultiCeps and Multi-frame Analysis:

Just like sound, optical echoes may occur in multinomial forms. Therefore, it is natural to question ceptrum's response to multiple visual echoes. Figure 3 shows multi-frame analysis of a simulated "accelerating" motion field - multiCeps.

Analytically, every pair of echoes generates a peak in the cepstrum plane, and these peaks in turn mutually reinforce one another. But the performance of cepstrum, although usually improves with multiple frames, is not dependent on the existence of

Figure 5: Trinocular image analysis. The cameras are in a right angle triangle, with the left camera directly underneath the top camera. The disparities between the left and right camera images is 12 pixels horizontally and between the top and bottom cameras 6 pixels vertically. The background is partially occluded by an object placed close to the cammera.

more than two frames. Moreover, unlike some spatio-temporal techniques, there is no assumption or requirement of constant linear velocity.

Figure 4 shows the use of cepstrum for stationary texture analysis. The peaks in cepstrum isolate the location of individual textons in the figure.

Finally, we examine cepstrum's performance for trinocular stereo analysis [DA91]. Figure 5 shows how cepstrum finds the disparity between the images generated by a trinocular stereo configured in a right angle triangle. As the result indicates there is only one horizontal, one vertical and one diagonal disparity present in this configuration; and multiCeps performed very well even in the presence of occluding objects and large disparities.

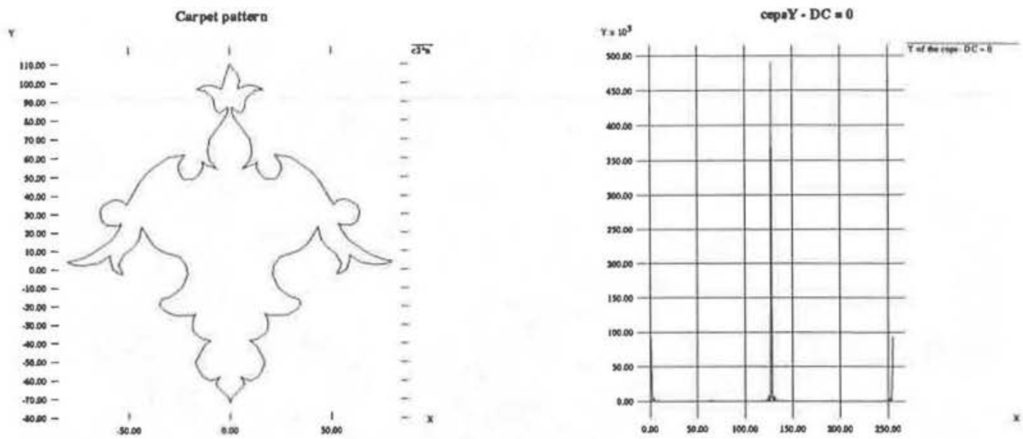


Figure 6: An axially symmetric curve, and the cepstrum of its parameterized Y coordinate. The second half of the curve was reflected and the DC component of the curve was made equal to zero.

5 Symmetric Boundary Analysis:

There are a number of psychophysical studies on the importance of boundary symmetries in human recognition, memory, and learning (for a survey see [Wil91b]). These have sparked interest in the use of symmetries in computational object recognition, specifically through perceptual organization [Wil91b, Low87]. A quick mathematical analysis reveals, however, that cepstrum – without preprocessing – is rather inadequate for detection of "single" axial symmetry. In short, the parameterized coordinates of a singularly symmetric curve are geometrically transformed even or odd functions; and the cepstrum of an even or odd function is always an even function. To overcome this problem, one has to first locate the axis of symmetry and simply reverse the second half of the curve. Figure 6 shows an axially symmetric figure (with the coordinates of its second half reflected); the cepstrum of its Y coordinate – after the DC component has been reduced to zero – shows a strong peak at the center.

It may seem simpler to check if the real or imaginary component of the Fourier transform of each parameter are zero functions. For numerical results, however, this will almost never be the case. Additionally, cepstrum provides a level of robustness toward noise which may prove desirable.

The main difficulty seems to be the sensitivity of cepstrum to the proper localization of the axis of symmetry. This is not a problem with multiply symmetric figures

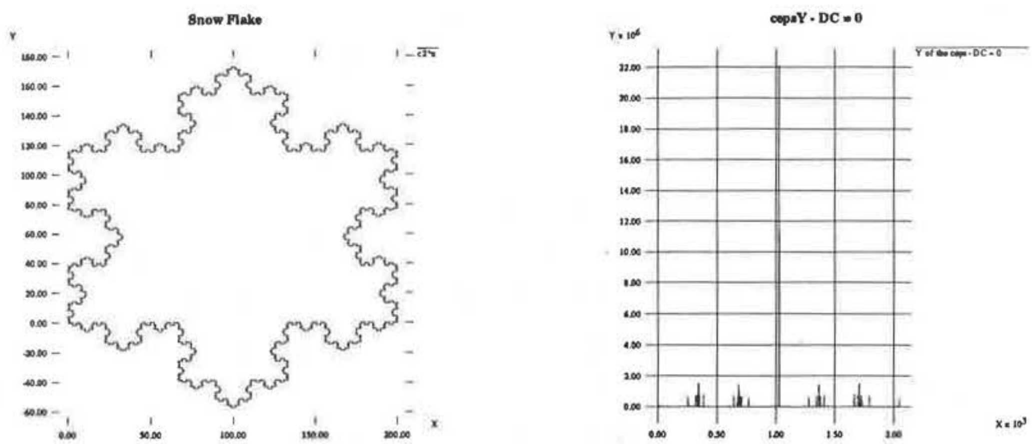


Figure 7: An axially symmetric curve, and the cepstrum for the new parameterized Y coordinate

however. In this case the parameterized coordinates of the curve consist of repeated even and odd functions which in turn become positive and negative echoes of one another. Consequently, even without finding the axis of symmetry, cepstrum should detect the symmetries of such boundaries. Figure 7 shows a multiply symmetric fractal curve. The small peaks are the result of the smaller symmetries.

It is noteworthy that multiple symmetries are also recognized faster by human observers than simple single symmetries. Moreover, it has been postulated that human observers also select potential axes of symmetries before processing boundaries [PH78].

6 Comparison to Standard Techniques:

In the previous sections, using the visual echo abstraction, we showed the use of cepstrum for a variety of visual tasks. In this section we compare cepstrum to other computational routines.

To date cepstrum routines have proved to be the most powerful technique for detection of echoes. Reddy and Rao, for instance, showed that differential cepstrum performed better than waveform delay analysis, or unwrapped phase averaging[RR87]. Such performance differences will become even more acute for images and higher dimensional signals, primarily due to greater possibility of errors in multidimensional phase unwrapping[Dud77]. Mitra and Lee [LMK89] showed that cepstrum is much more robust in the presence of noise than phase correlation [KH75].

For computational vision, visual motion analysis provides the richest area for the comparative analysis of cepstrum. Time varying image analysis plays a significant role in segmentation of scenes, encoding 3D information, egomotion estimation, object tracking, determination of focus of attention, and estimation of time to collision [Nak85]. Barron, Fleet, Beauchemin, and Burkitt[BFBB92] compared several motion algorithms, notably spatio-temporal filtering [Hee87], correlation matching techniques [Ana89], gradient based[HS81], and phase based approaches[FJ90]. We have not performed as detailed an analysis, but we have examined the performance of power cepstrum relative to phase correlation[KH75], gradient based approaches [HS81], and correlation [LBP88]. On synthetic motion images, compared to the first two techniques, we found cepstrum to perform very well particularly in the presence of noise. Correlation, on the other hand, is mathematically optimal in the presence of uncorrelated Gaussian noise, and performed strongly when used with synthetic images. Consequently, we explored the use of third order cumulants [NR87] [TG92] and Bicepstrum [NP88] for analysis of motions in one dimensional signals. These techniques are mathematically optimal in the presence of uncorrelated or "correlated" Gaussian noise. The main problem with these techniques is that they normally require large data sequences.

Analytically, cepstrum is obviously a "nonlinear" matching technique. Olson and Coombs [OC90] showed that cepstrum can be interpreted as an adaptive correlation mechanism. But, more significantly, the logarithm in cepstrum behaves as a smoothing filter by compressing the effect of narrow band signals, and emphasizing the broad band portions of the image (i.e., edges). This behavior is quite different from many linear techniques, specially the first and second gradient techniques.

Comparing techniques analytically, or even numerically, is a good academic exercise, and provides a valuable guide to improving the present methodologies. In fact, the development of cepsCorr was partially a consequence of comparisons that we drew between cepstrum and different correlative approaches [LBP88, Ana89, Fua91]. Generally, however, we found that drawing firm conclusions on such comparisons can be rather premature and deluding. Specifically, there are constant improvements introduced to traditional methods, which render such comparisons outdated (see for instance [CV92].

The next section tries to answer more specific questions about cepstrum — its strengths, and its weaknesses.

7 Frequently Asked Questions and Issues:

This section is intended to address some of the frequently asked questions about cepstrum, and to point out its strengths and limitations particularly as they relate to computational vision.

Sign Ambiguity in Disparity Calculation: Perhaps one of the main limitations of power cepstrum – which is overlooked by most researchers – is its sign ambiguity. To be more specific, because power cepstrum is a positive even function, a peak at location (u, v) may also be referring to a disparity $(-u, -v)$. Therefore, if the two patches used in Figure 1 are placed in the reverse order, the cepstrum outcome will be exactly the same. This shortcoming of power cepstrum is trivially overcome by a simple post processing technique (such as sum of the absolute differences) which easily eliminates the incorrect alternative.

Furthermore, it is noteworthy that our improvements to differential cepstrum (i.e. diffCepsSin) results in correct disparity calculations without any sign ambiguities or requirement for post processing [BL93].

Presence of Multiple Motions: A second frequently asked question about optical flow techniques is their behavior in the presence of multiple motions. Figure 8 shows two randomly generated patterns that have gone through arbitrary translations. The cepstrum of the resulting figure has two peaks corresponding to the correct disparities. This shows that cepstrum can be tuned to detect multiple moving objects, specular reflection, transparent surfaces and motion boundaries – a property not applicable to many other procedures.

False Peak at $(0, 0)$: From the above discussion, one can easily be lead to a simple technique for the figure/ground segmentation of motion or stereo sequences. Namely, if there is a strong second peak which corresponds to $(0, 0)$ disparity, it must have resulted from an occluding boundary. Even though often applicable, this approach may fail when analyzing smooth images. That is, in the presence of real disparity,

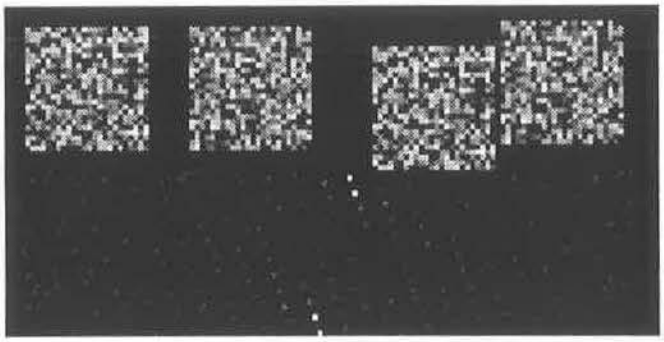


Figure 8: In the figure above two 24x24 random images were produced. One block was moved by 4 4 and the other by -1 -3. The Peaks in cepstrum show two peaks corresponding to the correct disparities.

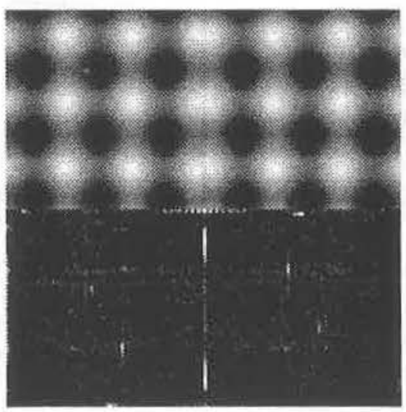


Figure 9: Power cepstrum of a moving plaid function

with no occluding boundaries present, cepstrum somtimes results in a strong primary or secondary peak at $(0,0)$ disparity. This behavior is explicable in terms of cepstrum's affinity to compress narrow band (i.e. smooth) signals.

Even though uncommon, this effect may result in incorrect interpretation of the image. It was this behavior that partially resulted in the development of cepsCorr.

Cepstrum as Deconvolution Filter: One of the main uses of cepstrum, and particularly complex cepstrum, is for deconvolution and deblurring purposes[LM91]. This is easily understood using the mathematics of cepstrum. The first Fourier transform turns a convolution into a multiplication, which in turn becomes an addition by the logarithm operation.

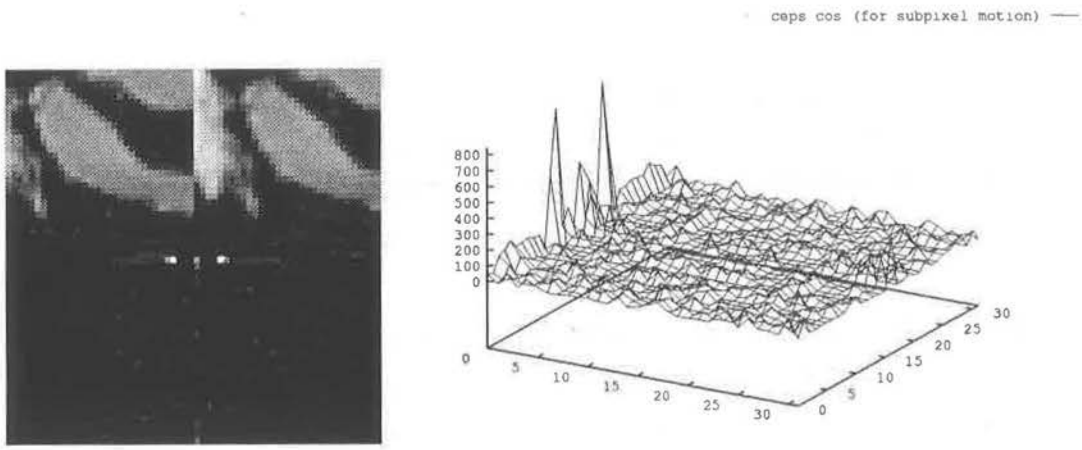
In fact, an echo is nothing more than the original signal convolved with a delayed delta; and it is this deconvolved delta that we use to measure the echo arrival period. The positive aspect of this behavior is that cepstrum is not adversely effected by effects such as motion blur, since it deconvolves and concenterates these effects at the boundaries of the resulted signal. On the other hand, this behavior does imply that patterns generated by the convolution mechanism are not easily analyzed by cepstrum. Figure 9 shows how the power cepstrum of a plaid function and its echo is decomposed into horizontal and vertical cosine functions.

The psychophysical implication of this phenomenon is that the ocular dominant columns in the primate visual cortex can not ustilize cepstrum as their processing mechanism[Wil91a].

Windowing Operations: One of the main questions that arise in any correlation based technique is the choice of window size used in matching. As Yeshurun and Schwartz [YS89] pointed out, the size of the cepstral window should be at least twice the size of the expected disparity field. An implication of this point is that cepstrum, like many linear filtering routines, is not ideal for measurement of large disparities caused by small objects. This effect was another motivation for the development of cepsCorr.

Another problem caused by windowing, arises from using a box filter to clip the regions of interest. This mechanism results in ringing effects that may cause artifacts interfering with the correct disparity peak. An obvious solution, both to improve the signal to noise ratio, and to emphasize the disparity measurement of the center pixels, is to multiply the matching windows with Gaussian or Weiner filters.

A word of caution, however, should be offered in here. Mathematically, multiplying the windows of Figure 1 by Gaussians is equivalent to convolving the Fourier of the concatenated images with a cosine Gabor function. This convolution can emphasize, or conversely deemphasize, the cepstral peak depending on the covariance of the Gabor.



(a) (b)

Figure 10: Positive cepsCos for a simulated subpixel motion. a) the figure and its positive cepsCos. b) the landscape of the cepstral plane.

Our numerical experiments show that if the Gaussian windows used in multiplication have small axial standard deviations, the cepstrum result will in fact be inferior.

Effects of Expansion and Rotation: Both Ludwig et.al. [LNN91] and Yeshurun [YS89] discussed the robustness of cepstrum with regards to expansion and rotation of images. Lee and Mitra[LMK89], on the other hand, used log polar mapping to transform rotation disparities into translation.

Briefly, cepstrum is robust with respect to rotations of up to 8 degrees and expansions of up to 6 percent. This becomes a valuable asset in analyzing looming and rotational motion, particularly when used in conjunction with cepsCorr.

Subpixel Motion: From the mathematics of section 2, it is obvious that for subpixel movement, the delta function representing the echo arrival period will be placed between two pixels, and hence the exact motion is recoverable by simple bilinear estimation. Olson and Coombs [OC90] reported that they verified this phenomenon experimentally. A word of caution, however, is applicable at this point. As pointed out earlier, traditional power cepstrum includes the cepstrum of the original signal as an additive noise. This additive noise can adversely effect the calculation of subpixel disparity.

Positive cepsCos on the other hand, reduces the false neighborhood values and

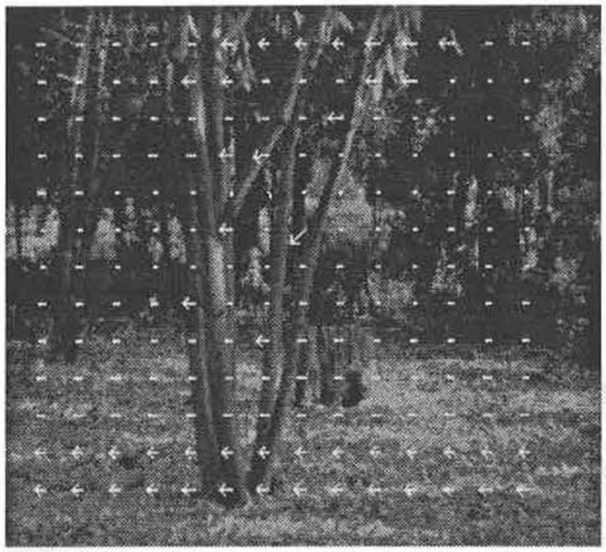


Figure 11: The optical flow field due to egomotion

increases the accuracy of subpixel disparity analysis. Figure 10 show the positive cepsCos for a simulated subpixel motion.

Effects of Noise: Hassab et. al. [HB76] provided a mathematically detailed treatment of cepstrum performance in the presence of uncorrelated Gaussian noise. Briefly stated, Gaussian noise modulates the cepstral peak and adds a reduced Gaussian noise to the cepstral result.

Parallel Processing: The computational complexity of cepstrum is equivalent to that of the Fourier transforms involved – $O(n \log n)$. Programming techniques [Ben82] or use of cepsCos in conjunction with Hartley transform reduce the computational cost of individual cepstrums. But one of the main properties of cepstrum is that it is parallelizable. Additionally, since cepstrum is a localized technique, it can be easily implemented on parallel hardware for parallel measurement of disparities.

We implemented cepstrum on a network of 24 transputer nodes. Figure 11 shows the result of this work and the behavior of cepstrum for analysis of a cluttered scene.

8 CepsCorr – A General Multi-evidential Correlative Approach to Detection:

The previous section provides an objective evaluation and criticism of cepstrum. In this section, we examine the significant cepstral defects, and provide a more general and robust solution.

When analyzing a translating object, in cepstrum and phase correlation, there is a portion of the two windows that does not overlap. This area is located at the boundaries of patches being analyzed and corresponds to the new regions of the object exposed, or the segments of the object that leave the window under consideration. The relative size of this area depends on the size of the window and the velocity of the moving object. The effects of this mismatch is, at best, increased added noise to the cepstral result, and at worst introduction of false peaks.

More significantly, like many optical flow techniques using fixed size static windows, a small fast moving object is never correctly identified. A small window will not contain most or any of the object, and a large window overshadows the motion by the effect of the background. Added to these are the effects of false $(0,0)$ peaks in the ceptrum and the aliasing noise in phase correlation.

To overcome these, and in fact utilize cepstrum's shortcomings, we introduced cepsCorr: a correlative approach where a window from from one image is matched against windows from a neighborhood in the second image. Essentially, cepsCorr replaces the sum of squared differences step used in correlation with a cepstrum. When in the vicinity of the proper disparity, individual cepstrums provide strong peaks – or evidences – that point to the correct motion field. These evidences can be combined using Baysian analysis [Sze88], relaxation labeling [HZ83], or multi-evidential reasoning [Tre68]. In our work so far, we have used a simplified approach, which involves choosing the motion with maximum response for the $(0,0)$ motion. What we found surprising was the strong difference in the magnitude of the cepstral peaks, and hence the large signal to noise ratio generated by cepsCorr. Even with small correlative windows on rather ambiguous regions cepsCorr's signal to noise ratio (S/N) was greater than 25. In the majority of our numerical experiments, cepsCorr's S/N were above two orders of magnitude.

Figure 12 shows cepsCorr's results for a complicated natural scene. The lower

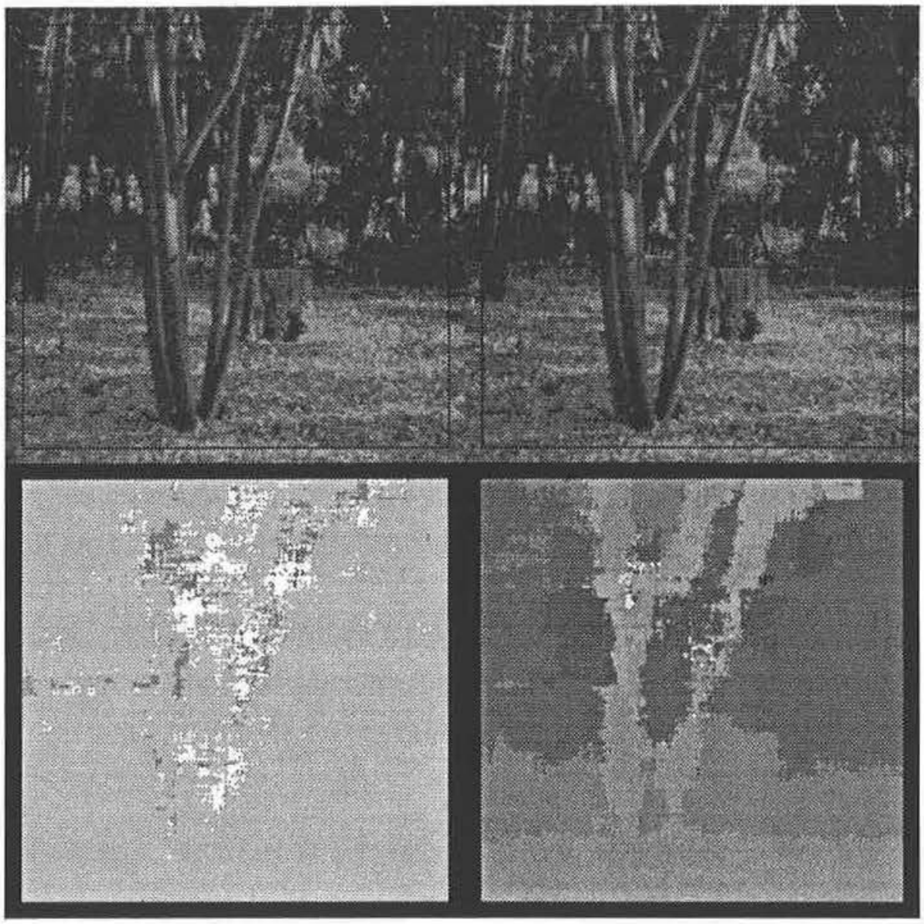


Figure 12: Horizontal and vertical disparity fields for a natural scene using cepsCorr. From knowledge of the direction of egomotion, the motion search region was restricted to the range of ± 1 in the vertical, and -2 to 7 in the horizontal direction.

left corner shows the vertical and the lower right displays the horizontal disparities calculated. Lighter shades depict larger displacements, and hence closer distances.

Obviously, complete correlation is not necessary if the initial peaks of the first fiew iterations agree on the correct disparity. We are presently examining several "voting" schemes to combine the result of multiple cepstrums. We are also investigating using different cepstrum techniques, along with cepsCos and phase correlation as the matching kernel for cepsCorr.

9 Summary and Conclusions:

This paper introduced the term "visual echoes" as a common framework for the analysis of multi-frame optical flow, binocular and trinocular stereo, stationary texture and boundary symmetries. We examined cepstral filtering, a powerful nonlinear adaptive technique for the retrieval of echoes, as a common methodology to address these visual routines. We then provided both computational and performance improvements to cepstrum which we hope will have positive impacts in other areas utilizing cepstrum.

Cepstrum is inherently robust in the presence of noise, rotation, expansion, and motion blur. We then examined cepstrum response to subpixel movement and multiple motions (such as transparent motion, or specular reflection), and put forth mathematically objective, but managable, analysis of cepstrum's weaknesses and strengths.

Finally we introduced multi-evidential correlation – cepsCorr – as a framework to improve disparity calculations by collecting multiple evidences and measurements while reducing the difference between the matching elements. CepsCorr provides a strong signal to noise ratio and a structure for detailed analysis of the signal's behavior by employing different matching routines as cepsCorr's kernel.

References

[AB85] E.H. Adelson and J.R. Bergen. Spatio-temporal energy models for the perception of motion. *J. OPT. SOC. AMER.*, pages 284–299, 1985.

[Ana89] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *Int. J. Comp. Vision*, 2:283–310, 1989.

[Ben82] J.L. Bentley. *Writing efficient programs.* Prentice Hall, Englewood Cliff, N.J., 1982.

[BFBB92] J.L. Barron, D.J. Fleet, S.S. Beauchemin, and T.A. Burkitt. Performance of optical flow techniques. In *Proc. CVPR*, pages 236–242, 1992.

[BHT62] B.P. Bogert, M.J.R. Healy, and J.W. Tukey. The quefrency analysis of time series for echoes: Cepstrum, psuedo-autocovariance, cross-cepstrum and saphe cracking. *Proc. of Symp. On Time Series Analysis*, 1962.

[BL92] E. Bandari and J.J. Little. Cepstral analysis of optical flow. Technical Report TR 92-6, University of British Columbia, 1992.

[BL93] E. Bandari and J.J. Little. Cepstral methods in computational vision. In *Nonlinear Image Processing IV; Proc. SPIE*, San Jose, CA, February 1993. SPIE, Bellingham, WA.

[CSK77] D.G. Childers, D.P. Skinner, and R.C. Kemerait. The cepstrum: a guide to processing. *Proc. IEEE*, 65(10), October 1977.

[CV92] M. Campani and A. Verri. Motion analysis from first order properties of optical flow. *CVGIP. Image Understanding*, 56(1):90–107, July 1992.

[DA91] U.R. Dhond and J.K. Aggarwal. A cost-benefit analysis of a third camera for stereo correspondence. *Int. J. Comp. Vision.*, 6(1):39–59, April 1991.

[Dud77] D.E. Dudgeon. The computation of two-dimensional cepstra. *IEEE Trans. Acoust., Speech and Signal Processing*, ASSP-25(6):476–484, December 1977.

[FJ90] D.J. Fleet and A.D. Jepson. Computation of component image velocity from local phase information. *Int. J. Comp. Vision*, 5:77–104, 1990.

[Fua91] Pascal Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. Technical Report 1369, INRIA, January 1991.

[HB76] J.C. Hassab and R. Boucher. A probabilisitc analysis of time delay extraction by the cepstrum in stationary gaussian noise. *IEEE Trans. Inf. Theory*, 22(4):444–53, 1976.

[Hee87] D. Heeger. Optical flow from spatio-temporal filters. *Proc. of 1st Int. Conf. Comput. Vision*, pages 181–190, 1987.

[HS81] B.K.P. Horn and B.G. Shunck. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.

[HZ83] R. A. Hummerl and S. W. Zucker. On the foundations of relaxation labeling processes. *IEEE Trans. Patt. Analysis. Mach. Intelligence*, 5:267–287, 1983.

[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] J.J. Little, H.H. Bulthoff, and T. Poggio. Parallel optical flow using local voting. *Proc. Int. Conf. Comp. Vision.*, pages 454–459, December 1988.

[LKM87] D.J. Lee, T.F. Krile, and S. Mitra. Digital registeration technique for sequential images. *Proc. Photo-Opt. Instrum. Eng.*, pages 293–300, 1987.

[LM91] T.R. Lewis and S. Mitra. Application of a blind-deconvolution restoration technique to space imagery. *Adaptive Signal Processing; Proc. SPIE*, pages 221–6, 1991.

[LMK89] D.J. Lee, S. Mitra, and T.F. Krile. Analysis of sequential complex images, using feature extraction and two-dimensioanl cepstrum techniques. *J. Opt. Soc. Am.*, 6(6), 1989.

[LNN91] K.O. Ludwig, H. Neumann, and B. Neumann. Local stereoscopic depth estimation using ocular stripe maps. *Proc. European Conf. Comp. Vision*, pages 373–377, 1991.

[Low87] David G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31:355–395, 1987.

[Nak85] K. Nakayama. Biological motion processing: a review. *Vision Research*, 25:625–660, 1985.

[NP88] C.L. Nikias and R. Pan. Time delay estimation in unknown gaussian spatially correlated noise. *IEEE Trans. Acoust., Speech, and Signal Processing*, 36(11), November 1988.

[NR87] L.C. Nikias and M.R. Raghuveer. Bispectrum estimation: a digital signal processing framework. *Proc. IEEE*, 75(7):869–92, July 1987.

[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*, 56:21364–1291, 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.

[PH78] S.E. Palmer and K. Hemengeway. Orientation and symmetry: Effects of multiple, rotational, and near symmetries. *J. of Exp. Psych.: Human Perc. and Performance*, 4(4):691–702, 1978.

[RR87] G.R. Reddy and V.V. Rao. Signal delay and waveform estimation through differential cepstrum averaging. *IEEE Trans. Acoust., Speech, and Signal Processing*, 35(10):1487–89, October 1987.

[Sze88] R. S. Szeliski. Bayesian modeling of uncertainty in low-level vision, 1988.

[TG92] M.K. Tsatsanis and B.G. Giannakis. Object and texture classification using higher order statistics. *IEEE Trans. Patt. Analysis. Mach. Intelligence*, 14(7):733–50, July 1992.

[Tre68] H.L. Van Trees. *Detection, estimation, and modulation theory.* Wieley, N.Y., N.Y., 1968.

[Wil91a] H. Wilson. personal communication, 1991. Spatial Vision.

[Wil91b] S. E. Wilson. Perceptual organization and symmetry in visual object recognition. *Master's Thesis, University of British Columbia*, 1991.

[WY91] F. Wang and P. Yip. Cepstrum analysis using discrete trigonometric transforms. *IEEE Trans. Acoust., Speech, and Signal Processing*, 39(2):538–541, February 1991.

[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.