DECEMBER 2001
VOL. 12, NO. 1

SPIE's International Technical Group Newsletter

ELECTRONIC IMAGING

Autowaves for image processing

Advances in synergetics and in the understanding of the mammalian retina and visual cortex have lead to new approaches to processing visual information. Patterns of traveling waves of neural activation have been found in both retina and in cortex. Autowaves represent a particular class of these spatio-temporal patterns, which propagate in an active media (i.e. neural network) at the expense of the energy stored in the medium.[1] They have some typical characteristics that are fundamentally different from those of classical waves in conservative systems. Autowaves do not reflect from inhomogeneities: there is no interference because two colliding autowaves annihilate each other. Nonetheless, both autowaves and classical waves share the property of diffraction. These properties provide invariance for image processing under translation, rotation, and scaling, that make autowaves useful for image processing. Their exploitation, using models of reaction-diffusion systems implemented using the cellular neural network architecture, provides an opportunity for us to develop novel and efficient spatio-temporal image processing techniques.[2–5]

Our autowave-based image processing approach uses a pulse-coupled neural network (PCNN). The PCNN is, to a very large extent, based on the Eckhorn model of the cat visual cortex.[6] The basic simplified structure of the PCNN processor for a 2-D input image is shown in Figure 1. An input gray-scale image is represented as an array of M×N normalized intensity values. Then the array is fed in at the M×N inputs of the PCNN. The network finally processes the array to produce a series of binary images containing the segmentation result.

The processing is implemented in the following way. First, the dynamic threshold of each neural node significantly increases when the neuron fires, producing a binary impulse, then the threshold value decays. When the threshold falls below the respective neuron's potential, the neuron fires again, which raises the threshold again. The process continues creating binary pulses for each neuron. While this process is underway, neurons encourage their neighbors to fire simultaneously because of the strength of the excitatory connections between them. The firing neurons begin to communicate with their nearest neighbors that, in turn, communicate with their neighbors. The result is a traveling activation wave that expands from active regions. Thus, if a group of neurons is close to firing, then one neuron can trigger the whole group.

As a result of this linking between neurons, the pulsing activity of those invoked leads to synchronization between groups of neurons corresponding to sub-regions of the image with similar properties, and produces temporal series of binary images. These phenomena of synchronization and traveling waves support image processing such as image noise removal, segmentation, edge extraction, skeletonization, and motion detection.

Real images are noisy. Preliminary image smoothing,

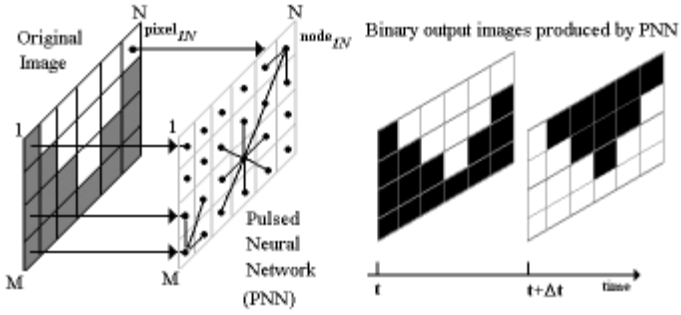


Figure 1. Framework of image processing using a pulsed cellular neural network. In the PCNN, output "white" pixels represent firing neural nodes, "black" pixels represent silent nodes.

continued on p. 8

Calendar *— See page 10*

Technical Group Registration Form *— See page 11*

NEWSLETTER NOW AVAILABLE ON-LINE

Technical Group members are being offered the option of receiving the Electronic Imaging Newsletter in an electronic format. An e-mail notice is being sent to all group members advising you of the web site location for this issue and asking you to choose between the electronic or printed version for future issues. If you have not yet received this e-mail message, then SPIE does not have your correct e-mail address in our database. To receive future issues of this newsletter in the electronic format please send your e-mail address to spie-membership@spie.org with the word EI in the subject line of the message and the words "Electronic version" in the body of the message.

If you prefer to continue to receive the newsletter in the printed format, but want to send your correct e-mail address for our database, include the words "Print version preferred" in the body of your message.

Change detection in aerial stereo pairs at different dates

Change detection in aerial imagery is an important task for many applications (cartographic, agricultural, military, etc.). A review of existing techniques can be found in Reference 1. The aim of this application is to detect changes in an aerial scene by comparing stereo pairs taken at intervals of several years in order to update a database. Figure 1 gives some examples of changes that can be found in a given area.

Figure 1. Example of a scene taken at five year intervals. We can see changes in different classes of objects inside this scene (vegetation, buildings, water areas, etc.).

The result of the proposed algorithm is a set of cartographic locations that have a high likelihood of containing changes. Each location will be submitted to a human operator who will either validate the given change and update the database, or reject it. We are mainly interested in changes occurring for a specific class of objects: buildings. To isolate new constructions, we provide an algorithm that works in two steps.

Depth comparison

First, during a focusing phase, we aim to eliminate a large part of the scene without losing any actual changes. This is achieved through a comparison of the Digital Elevation Model (DEM) between the two different dates. We used the depth map computation algorithm described in Reference 2. Other depth algorithms and references can be found in Reference 4. The old depth map is computed with the old stereo pair, and the new depth map is computed with the new one.

Here, median values of the depth histograms for small regions (5×5m) at the two different dates are compared. The DEM comparison leads to focusing areas. Each focusing area is a set of four images: a stereo pair of the area at the old date and a stereo pair at the new date. The technical details of this step can be found in Reference 3. Some of these regions of interest (ROI) contain true changes and some no relevant changes at all. The true changes have to be separated from the false alarms.

Region of interest classification

In the second phase, regions of interest are classified. Each ROI is described by four images: stereo pairs of the focusing area at the first and second dates. To decide whether the ROI contains a change or not, each of the four images is classified as "building" or "non-building". The building versus non-building classifier is a combination of several decision trees induced by a learning stage. Each node of a decision tree is identified with a graph of features, which is more likely to describe buildings than background. Figure 2 gives an example of a decision tree and the feature graphs associated at each node.

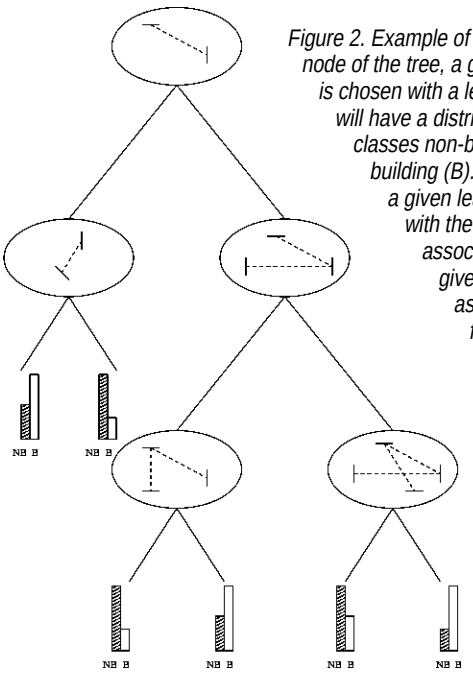


Figure 2. Example of a decision tree. At each node of the tree, a graph is associated that is chosen with a learning set. Each leaf will have a distribution of our two classes non-building (NB) and building (B). An example arriving at a given leaf will be associated with the mode of the distribution associated to this leaf. At a given internal node, as soon as the associated graph is found on the image, we will go to the right subtree.

Finally, the classification results at the two different dates are compared. An area is selected as soon as one of the images from the old stereo pair is classified as background, and at least one of the images from the new stereo pair is classified as building. Again, the technical details for this classification step can be found in Reference 3.

Results

Figure 3 illustrates some result. Our approach has been tested on two different scenes: each contains between 150 and 200 changes. The false negative rate is close to 10% of the overall number of changes in the scene. In order to have an idea of the accuracy of our algorithm, we will have to make further tests on different scenes. In the near future we will have to face change detection problems for different resolutions. We think that our algorithm is well designed for resolutions between 40cm and 1m.

Limitations and future work

Many false-negative examples are due to artifacts and errors made by our DEM. We will have to improve our DEM algorithm for the change detection task: a good DEM must furnish reliable depth information even if this information is sparse.

We will also have to use a vegetation detector in order to reduce the number of false positives.

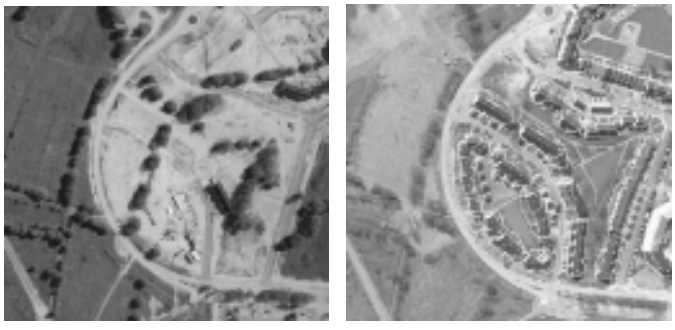


Figure 3. Example of a change detection. The image on the left shows an area at a given date t_2 and the image on the right shows the same area at a date t_2 with $t_2 > t_1$ with the result of our algorithm superimposed.

Vegetation areas are challenging for both the depth map algorithm and the classification algorithm. Inside vegetation areas, we will have a lot of random depth variations, and a lot of edges in almost every direction: hard to deal with using our graph-matching algorithm.

Some complementary experiments will be achieved, with several different human operators, in order to compare a fully-manual database updating approach with ours. With these experiments, we will be able to see what input a human

continued on p. 8

Image restoration: beyond wavelets

Have you ever been confronted with the complicated situation of trying to get something out of noisy data? Experts would say that wavelet techniques are the ideal tool for such a task, and it is true that wavelets and related multiscale representations pervade all areas of signal processing. The recent inclusion of wavelet algorithms in JPEG 2000—the new still-picture compression standard—testifies to the lasting and significant impact of wavelets. The reason for their success is the fact that the wavelets basis represents a large class of signals well, and therefore allows us to detect roughly isotropic elements occurring at all spatial scales and locations.

As noise in the physical sciences is often not Gaussian, the modelling of many kinds of noise in wavelet space (Poisson noise, a combination of Gaussian and Poisson noise, non-stationary noise, etc. ...) has been a key step for the use of wavelets in scientific, medical, or industrial applications.[1] Extensive wavelet packages exist now, both commercial[2] and non commercial,[3] that allow any researcher, doctor, or engineer to analyze his or her data using wavelets.

Figure 1 shows the result after applying the wavelet-filtering method to a real spectrum. Figure 1, bottom left, shows the difference between the original and filtered spectrum (residual). As we can see, the residual contains only noise. Note how the important spectral lines are accurately preserved.

After such results, is the noise removal problem definitively solved? Not exactly: for a 2D or 3D data set, the wavelet basis presents some limitations because it is not adapted to the detection of highly anisotropic elements, such lines in a image, or sheets in cube. Recently, other multiscale systems like curvelets[4] and ridgelets,[5] which are very different from wavelet-like systems, have been developed. Curvelets and ridgelets take the form of basis elements that exhibit very high directional sensitivity and are highly anisotropic. A digital implementation of both the ridgelet and the curvelet transform for image noise removal has been described in Reference 6.

To understand the main difference between wavelets and ridgelets, consider an image that contains a vertical band embedded in relatively-large-amplitude white noise. Figure 2 shows such an image. Note that it is not possible to distinguish the band by eye. The wavelet transform (undecimated wavelet transform) is also incapable of detecting the presence of this object, while the ridgelet transform detects it clearly (above the 5s detection level). Another example is presented in Figure 2, which shows a part of the Saturn rings. The wavelet filtering is clearly not as good as the curvelet filtering, which better respects the anisotropic features contained in the data.

Although the results obtained by simply thresholding the curvelet expansion are encouraging, there is, of course, ample room for improvement. Indeed, each transform has its own area of expertise and this complementarity may be of great potential. In Starck et al.,[7] a combined filtering method has been proposed which allows us to build a solution, ensuring that it incorporates information judged as significant by *any* of our representations. The idea is to force the solution to be smoothed, but under the constraint that any significant coefficient (i.e. Fourier, wavelet, ..., ridgelet or curvelet coefficient larger than a given detection level) detected by a given trans-

continued on p. 9

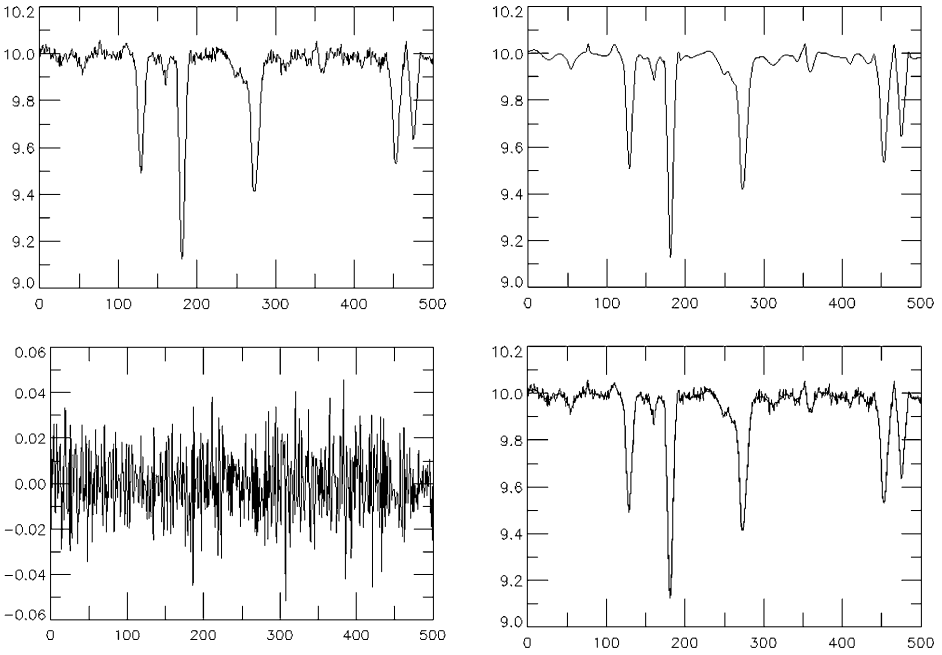


Figure 1. Top: real spectrum and filtered spectrum. Bottom: both noisy and filtered spectra overplotted, and difference between the spectrum and the filtered data. As we can see, the residual contains only noise.

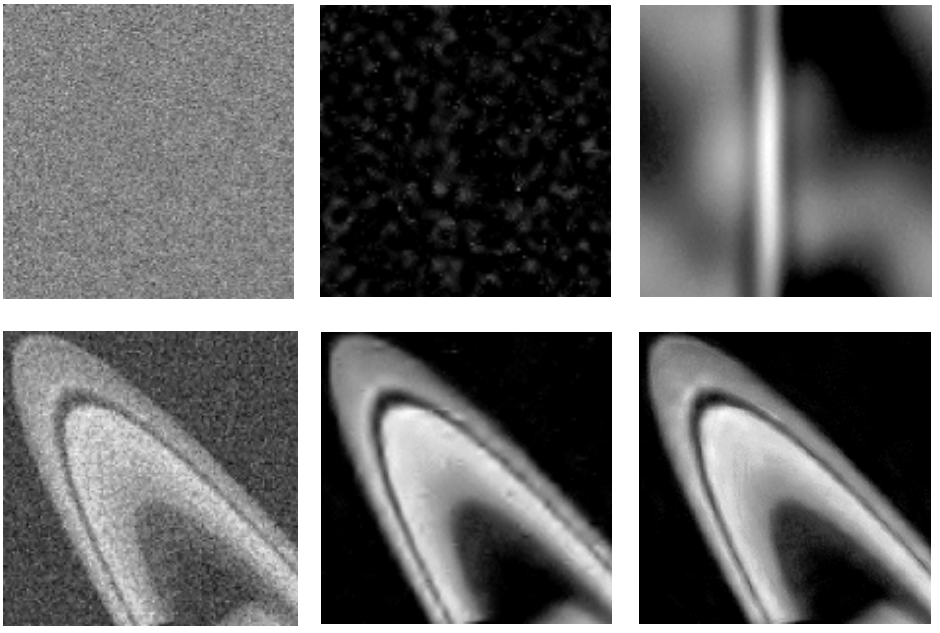


Figure 2. Top, respectively from left to right: original image containing a vertical band embedded in white noise, reconstructed image using the undecimated wavelet transform, and reconstructed image using the ridgelet transform. Bottom, respectively from left to right: noisy data (Saturn rings), reconstructed image for the wavelet coefficients, and reconstructed image from the curvelet coefficients.

Text detection in binarized images of advertisements

Documents with complex and arbitrary structures are becoming more and more widespread thanks to rapid advances in publishing technology. Advertisements published in magazines or on the Web, which are rich in color and texture and contain other special effects to attract readers, are examples of this. Many modern office systems using commercial character recognition systems can recognize color text only with strict constraints imposed on the quality of the document image. This limits the areas where such systems can be applied to simple documents, like technical journals, the structure of which are well-defined and constrained.

The main reason for the failure of character recognition seems to be an inability to correctly identify text regions in an image. Our team is trying to tackle this text detection task. Research started from a search for simple features relying on edge information extracted from grayscale advertisement images.[1] Grayscale images were used instead of color because it was known that most of the edge information is contained in the luminance component of a color image. This attempt led to promising results, but the false positive rate was also significant. It became obvious that deeper analysis—involving binary, grayscale, and color information—was needed for accurate and reliable text detection. Since binary images are simpler to analyze, research on text detection started with these images.

Each original color image taken from a database was binarized with several global thresholds. An example of one of the binary images obtained is shown in Figure 1(a), where both black text on a white background, and white text on a black background are present. In addition, the font sizes significantly vary within the image and vertically-oriented text is included. (We worked on the expectation that text could be aligned either horizontally or vertically, which is most commonly the case.)

To cope with the conditions mentioned, we proposed a new method for text detection. Although it relies on a set of heuristically-chosen parameters, it has a number of advantages over existing methods. Firstly, the number of tunable parameters is moderately small, and each of them is typically limited by a narrow interval of values. Secondly, definitions of well-known properties of text characters are revised to allow extra flexibility when dealing with characters of various font styles and sizes.

The method begins with nonlinear order-statistic filtering, which eliminates small isolated

Figure 1. Text detection from binary images: (a) binary image; (b) text detection results.

configurations of black and white pixels. Connected component analysis, assuming that text is black, determines the parameters of the bounding boxes of the components. Knowing the first and last rows of each bounding box helps to identify components composing initial line candidates. Since these may span two text columns, line partitioning is done based on regular distances between characters belonging to the same text line. At the same time, candidates that are too short are dismissed from further analysis. Usually these come from pictures, so this operation reduces the false-positive rate. However, lines that are too short may also include characters. Hence, components resembling characters are searched for both within and outside (in the immediate vicinity) of each line candidate. This search is based on the fact that, given a fixed font style and size, characters—independently of a script or a language—are composed of strokes of approximately constant width. As a result, a set of horizontal lines containing black text is generated.

The same operations as just described are applied to the remaining components in order to detect vertical lines of black text. The only exception is that height-related features are replaced with width-related ones. White text on a black background is then located by repeating all steps required to detect black text.

Text detection results are given in Figure 1(b), where black and white detected text is copied as it is from the image in Figure 1(a), while non-text is presented in gray. Both black and white text was successfully extracted independently on font size and text orientation. Relatively small portions of pictures were identified as text, too, but it was expected that non-text that looks like text would be indistinguishable from the real thing.

Though the results obtained are very promising, binary data alone cannot reliably detect text in all cases. For example, characters can go undetected if they touch pictures. Future research will therefore concentrate on utilizing grayscale and color features to alleviate such problems.

Oleg Okun and Matti Pietikäinen
Machine Vision and Intelligent Systems Group
Infotech Oulu and
Department of Electrical Engineering
University of Oulu
P.O.Box 4500
FIN-90014, Finland
Email: {oleg, mkp}@ee.oulu.fi
URL: http://www.ee.oulu.fi/research/imag/document/

References

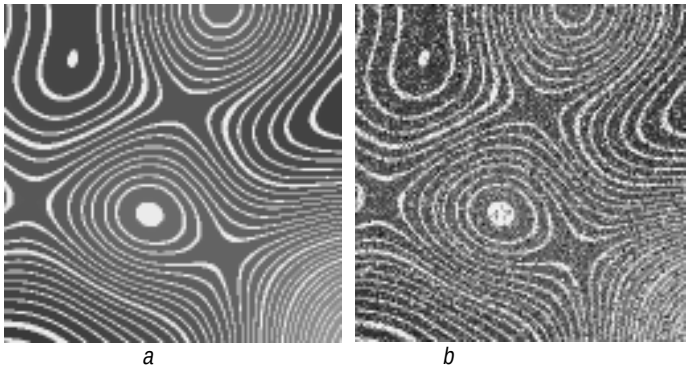
1. M. Pietikäinen and O. Okun, Edge-based method for text detection from complex document images, **Proc. 6th Int'l Conf. on Document Analysis and Recognition**, Seattle, WA, USA, pp. 286-291, 10-13 September 2001.

Noise suppression using nonlinear filters with spatially connected neighborhoods

Many nonlinear filters incorporating rank-order operations have been introduced in computer and optical research in recent years. There are several different classes of such filters: median filters, multistage- and multilevel-median filters, stack filters, order-statistics filters, morphological filters, and rank-order filters. These have all proven to be very effective for the removal of additive and impulsive noise, and for enhancing and restoring images. Moreover, they exhibit excellent robustness properties and can provide solutions in many cases where linear filters are inappropriate. The primary reason for their success in image processing is that they can suppress noise without destroying important image details such as edges and fine lines.

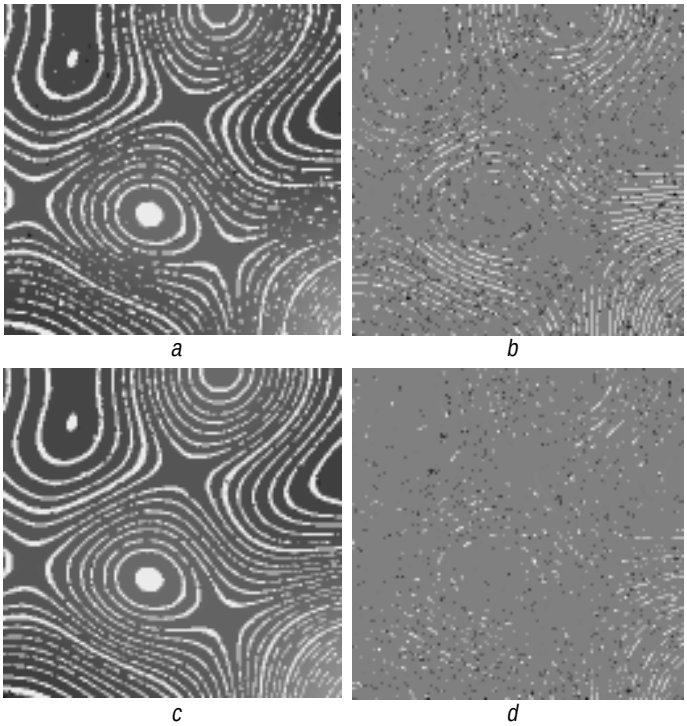
In the design of rank-order filters, image elements of a moving window are sorted in ascending order called the variational row. The output of a rank-order filter is a function over the elements in the variational row around the central element of the window. Since rank-order filters take into account local image content (local statistics), the rank-order filtering is locally adaptive. A drawback of conventional rank-order filters is that they only weakly exploit spatial relations between image elements, because they perform the re-ordering of elements of a 2D moving window into a 1D sequence (variational row).

We suggest a new class of rank-order filters that explicitly use spatial relations between image elements.[1] To produce the output, the filters use spatial and rank information from spatially connected areas of the input image within a moving window. We use the notion of neighborhood to define various useful structures in the image. The three types of neighborhood are defined as follows. The CEV-neighborhood is the subset of pixels from the moving window that are spatially connected with the central pixel, and whose values deviate from the value of the central pixel at most by predetermined quantities. The CKNV-neighborhood is a subset of a specified number K pixels from the moving window that are spatially connected with the central pixel, and whose values are nearest to the value of the central pixel. Finally, the CER-neighborhood is the subset of pixels from the moving window that are spatially connected with the central pixel, and whose ranks deviate from that of the central pixel at most by predetermined quantities.



a b

Figure 1. (a) Test image. (b) Noisy image (impulsive noise).



a b

c d

Figure 2. (a) Median filter. (b) Enhanced difference between original image and image processed with a median filter. (c) Our proposed filter. (d) The enhanced difference between the original image and that processed using our filter.

The choice of neighborhood is defined by the available *a priori* information on the processed image. If *a priori* information about the geometrical size K of the details to be preserved is known, then the CKNV-neighborhood can be used. The parameter K is chosen to be of the order of the detail area to be preserved after further processing. The choice of the CEV-neighborhood helps us to take into account *a priori* information about either the spread of the signal to be preserved, or the noise fluctuation to be suppressed. Finally, the CER-neighborhood is often used in edge-extraction algorithms and for suppression of noise with a distribution having heavy tails.

The output of the filtering is a value computed as a basic operation (sample mean, median value, minimum and maximum values) over the neighborhoods. The spatially connected neighborhoods are not formed across region boundaries, so noise suppression will not blur image edges as often happens with other techniques. Signal processing of an image that has been degraded due to impulsive noise is of interest in a variety of tasks. Computer experiments are carried out to illustrate the performance of a median filter of size 3×3 elements and the proposed algorithm. The proposed algorithm results either in the median value over the CEV-neighborhood, if the size of the neighborhood greater than a predetermined threshold, or in the median value over elements surrounding the central element. The size of the moving window is 5×5 elements. Figures 1(a) and (b) show a test interferogram image containing fine lines and the same image corrupted with impulsive noise. The probability of an impulse occurring is 0.2, and if it occurs it can be positive or negative with equal probability. Figures 2 (a) and (c) show the processed images with the median filter and the proposed algorithm, respectively. Figures 2 (b)(d) show the enhanced difference of the original image with the median filtered image, and with the filtered image with using the algorithm described here, respectively. It can clearly be seen that our algorithm significantly outperforms the conventional filter.

Vitaly Kober, Mikhail Mozerov, and Josué Alvarez-Borrego
CICESE, Departamento de Óptica
División de Física Aplicada
Km 107 Carretera Tijuana-Ensenada
Ensenada 22860, B.C., México
E-mail: vkober@cicese.mx

References

1. V. Kober, M. Mozerov, and J. Alvarez-Borrego, *Nonlinear filters with spatially-connected neighborhoods*, **Optical Engineering 40** (6), June 2001, pp. 971983.

Block-based MRF image modeling

Markov random field (MRF) models have been widely used for image segmentation and restoration problems. The beauty of MRF modeling is its ability to describe a large number of spatial interaction phenomena by means of local characteristics. Such models are typically used as priors to represent the continuity and the smoothness of the pixel-based spatial information. Conventional pixel-based MRF modeling, however, has a limited ability to describe large-scale behavior. The model can be improved by adopting a larger neighborhood system, but this immediately increases the number of parameters to be estimated, which, in turn, increases the computational complexity significantly.

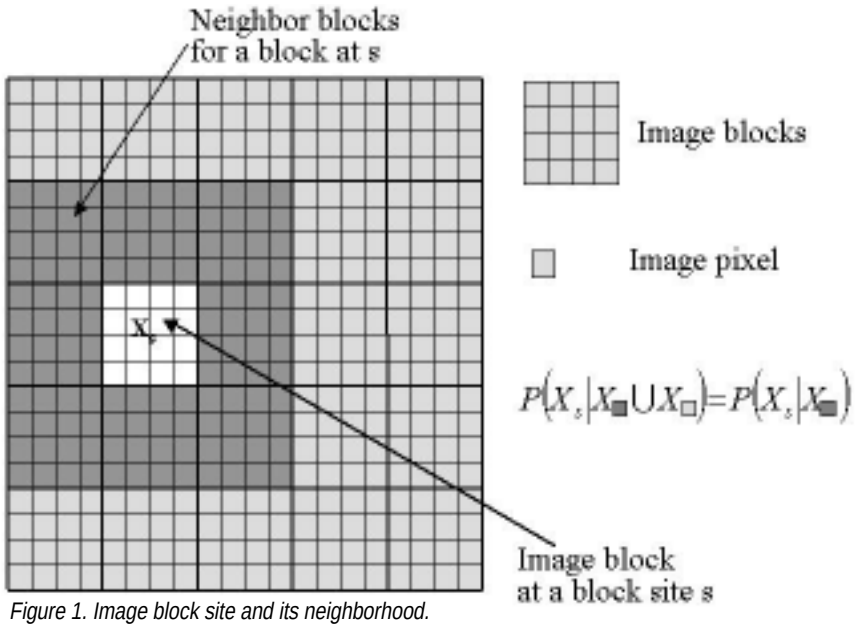


Figure 1. Image block site and its neighborhood.

To overcome this limitation, multi-resolution (or hierarchical) extensions of the MRF model have been considered. By decomposing the image data into different frequency components and scales, the MRF models can be used to describe the interactions between consecutive transition levels as well as adjacent sites at each resolution.

However, recalling that the merit of the MRF model is its ability to describe the global behavior of the image by repeated local updates, the role of the MRF model can be limited in a the multiresolution structure. That is, the coarse-to-fine (i.e., global-to-local) treatment of the image in the multiresolution environment may weaken the effect of the MRF modeling. In fact, it has been shown that the aforementioned limitation of the MRF models is not caused by the nature of MRF, but by inappropriate choice of local features and statistics.[1] For example, it has been experimentally shown that macro textures and shape patterns can arise from simple local features such as Gabor filters and Gestalt grouping rules for edge fragments. This implies that it is still possible to describe the large-scale behavior of the image by the conventional setting with a single resolution structure.

A plausible way to use the local features and statistics is to adopt block-based MRF modeling. Specifically, by dividing the image space into small, non-overlapping image blocks, and assigning random variables to each one, we can express local features, such as edges and texture, that are more helpful for the large-scale description. For example, we can assign two random variables to each image block. One of them is responsible for the continuity of the block label and the other represents the coloring of the image block given the underlying block label. Then, the set of all random variables of the image blocks constitutes the random field X for the hidden block label configuration, and the random field Y for the observable image block features. As shown in Figure 1, the random field X is assumed to be an MRF, i.e., the conditional probability of a random variable X_s at a block site s depends only on the block labels of its neighbor blocks. Given a realization y of Y, we can find an optimal MAP (maximum a posteriori) realization x^* of X that maximizes $P(X|Y)$ for all possible realizations of X. The global maximization of $P(X|Y)$ can be achieved through local updates of the associated MRF model.

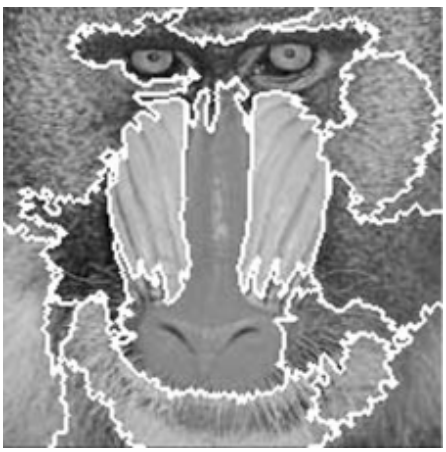


Figure 2. Example of segmented image.

In the application of block-based MRF image modeling to the image segmentation problem, the realization of the random variable for the block label is assumed to take one of six values.[2] The realization characterizes the image block into one of the following six classes: monotonic intensity block, texture (or non-directional edge) block, horizontal block, vertical block, and two diagonal blocks. Adopting the MAP criterion, the optimal block label configuration x^* yields the classification of the image blocks among one of the six block configurations. Note that the monotonic intensity blocks and the texture (or non-directional edge) blocks belong to interior homogeneous regions. So, they are clustered and are given region labels. Then, all remaining unlabelled blocks, including edge blocks, are assigned to the adjacent homogeneous regions. During this process, we can divide each unlabelled image block into four small blocks to increase the segmentation accuracy in the region boundary. We can repeatedly divide the unlabelled image blocks until we have a pixel level segmentation. As shown in Figure 2, the segmentation result yields large segments, which indicates that the block-based model can extract the large-scale behavior of the image.

Chee Sun Won
Dept. of Electronics Engineering
Dongguk University
Seoul, 100-715, Korea
E-mail: cswon@dongguk.edu

References

1. S. C. Zhu and Y. N. Wu, *From local features to global perception – A perspective of Gestalt psychology from Markov random field theory,* **Neurocomputing 26-27,** pp.939-945, 1999.
2. C. S. Won, *Block-based unsupervised natural image segmentation,* **Optical Engineering,** 39 (12), pp.3146-3153, December 2000.

New distortion measures for compression of remote-sensing data

Remote sensing applications for command and control, intelligence gathering, telemedical monitoring and other applications can require huge amounts of signal and image data to be sent over the information infrastructure. In many of these applications, the image data transmitted is not intended for viewing but rather is used to make automated decisions and inferences after being fused with other data at a site far from where they were recorded, and to accomplish this in a timely manner requires effective data compression schemes. In such applications, however, standard multimedia compression techniques are generally not highly effective. A key difference between compression for these applications and for multimedia is that the data is not generally consumed directly by humans, but rather is used to make decisions about, and estimations of, physical parameters reflected in the data. We are currently developing compression techniques for remote sensing applications that are specially designed for the purpose of optimizing the performance of subsequent estimation processing, and have shown that they outperform methods designed using more standard approaches to data compression. In such applications, it is crucial that the compression methods minimize the impact on the estimation performance, rather than stressing minimization of mean-square error (MSE) as is common in many compression techniques.

Important work has been done towards establishing some theoretical bounds, as well as some general theoretical underpinnings of estimation and decision using compressed data (e.g. Zhang and Berger). However, just as the important gains in compression for multimedia are being made by carefully exploiting the interaction of specific signal characteristics with specific consumption characteristics (e.g., psychology of vision, etc.), rather than just applying general theoretical results, major gains in compression for remote sensing can come from understanding how to exploit the interaction of specific signal characteristics and the parameters to be measured.

Some of the key issues that we are considering follow. First is the development of new distortion measures that provide an understanding of how to design algorithms for remote sensing applications: such methods should be general enough to allow easy generalization and extension to a broad array of applications, yet provide a means for exploiting application-specific characteristics. Second, we consider how to make trade-offs in the case where a user is interested in estimating multiple parameters but their requirements for compression are in conflict. Third, we are trying to understand how to compress remotely- collected data when it may have multiple uses that have conflicting compression requirements. Finally, we want to characterize the interaction between the compression technique and the communication network. Here we report on our efforts concerning the first of these issues.

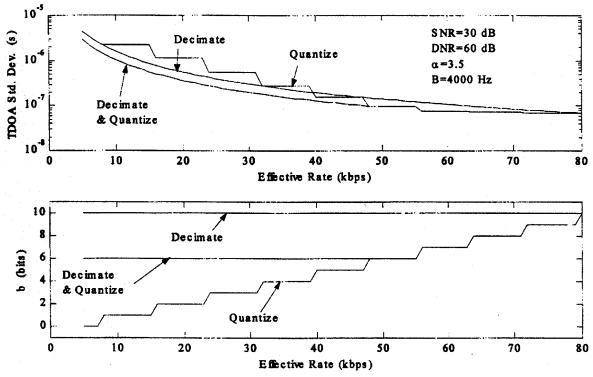


Figure 1. Rate-TDOA accuracy results for using decimation-only, quantization-only, and combined quantization and decimation. The top plot shows the TDOA accuracy versus the rate, and the bottom plot shows the number of bits used per sample at each rate for each case. It is clear that the combined approach is better for rates below 48kbps: at rates above 48kbps, the rate is high enough that the full signal bandwidth should be used according to the optimization rule we have derived. Therefore, the combined approach is most useful at low rates.

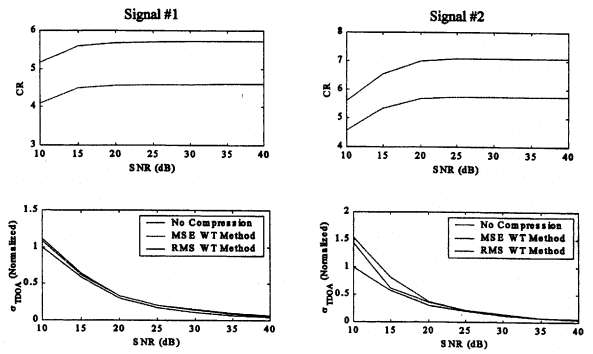


Figure 2. Compression results for using our sub-optimal non-MSE wavelet transform method compared to a pure MSE-based wavelet transform method for two different signals. Solid curves are for the no-compression case, the dashed curves are for the MSE-based case, and the dotted curves are for the sub-optimal RMS-measure based case. The top two plots show the compression ratio versus pre-compression signal SNR. The bottom two plots show TDOA accuracy versus pre-compression signal SNR. Note that the sub-optimal RMS-based method outperforms the MSE-based method.

The applications listed above, while by no means the only applications, are considered to be representative and important cases. Each one can be used as a vehicle to explore the issues we have described and to develop and test general methodologies. The problem that we are currently using as a development sandbox is the estimation of the shift between two 1D signals, and the extension to the corresponding 2D image problem, is also underway (the estimation of the translation between two different views of the same scene, perhaps for the purpose of alignment needed for subsequent automatic combination or comparison of the two images). The 1D problem of shift estimation is called time-difference-of-arrival (TDOA) estimation and arises in the remote-sensing problem of estimating the location of a source from signals received at multiple platforms.

Compression for shift estimation

To ensure maximum performance, it is necessary to employ a compression method that is designed specifically for this application. However, much of the effort in developing general lossy compression methods has focused on minimizing the mean-square error (MSE) due to compression: furthermore, even compression schemes developed for TDOA applications have limited their focus to minimizing the MSE. But when the goal is to estimate TDOA, the minimum MSE criterion is likely to fall short because it fails to exploit how the signal's structure impacts the parameter estimates. In particular, for remote-sensing estimation problems, the Cramer-Rao bound (CRB) provides guidance as to what signal characteristics are important.

Achieving significant compression gains for the emitter-location problem requires exploitation of how signal characteristics impact the TDOA accuracy. For example, it is known that the TDOA accuracy is inversely proportional to the RMS bandwidth of the signal's spectrum, called this because of its similarity to measuring the root-mean-square value of a probability density function. Thus, compression techniques that can significantly reduce the amount of data while negligibly impacting the signal's bandwidth have great potential. For instance, we recently obtained results that show that it is possible to exploit this idea through simple filtering and decimation to meet requirements on data transfer time that can't be met through quantization-only approaches designed to minimize MSE. Further, it has been shown that, for cases when minimum MSE approaches *can* meet the data-transfer time requirement, simple techniques that exploit the bandwidth characteris-

continues on p. 9

Autowaves for image processing

continued from front cover

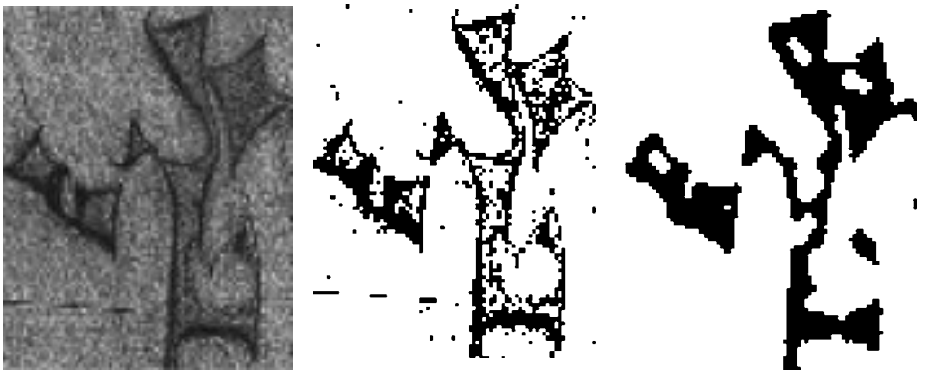


Figure 2. Image noise removal. (a) Original image. (b) Segmentation without noise removal. (c) Segmentation after noise removal.

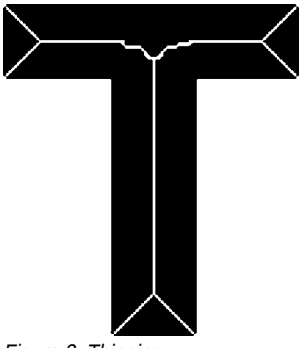


Figure 3. Thinning.

Figure 4. Traces left by traveling wavefronts from two objects moving in opposite directions at different velocities.

based on similarities of adjacent pixels, can support better segmentation. Some of the results from the application to segmentation of images of cross-sections for polymeric foam are shown in Figure 2 (a-c). As can be seen from the images in Figure 2(b and c), the combination of PCNN smoothing and PCNN segmentation produces segmentation that is less noisy. The process known as "fire-front" or "grass fire" propagation transform, which exploits the annihilation property of autowaves, is used for skeletonization. The object's skeleton is the locus of intersections of wavefronts propagated inwards with a constant speed from the border of the object (see Figure 3). The existence and interactions of autowaves also provide computational mechanisms to detect and characterize motion. Autowaves resulting from moving objects leave traces in the PCNN two-dimensional sheet of neural nodes. Traces left by wavefronts may be used for motion characterization (see Figure 4).

Further research on the combination of statistics extracted from wave propagation with statistical and classification techniques can facilitate development of algorithms for classification of images, recognition of objects in images, and motion characterization. In addition, wave-based processing is inherently parallel and can be exploited by advances in hardware, i.e. field programmable gate array technology (FPGA). FPGA implementation of image processing operations has the potential to achieve a speed-up of over two orders of magnitude compared to software implementations. This will be an area of future research for us, as well as focussing on wave-based motion detection and characterization.

Alexei N. Skourikhine
Safeguards Systems Group, MS T005,
Los Alamos National Laboratory,
Los Alamos, NM 87545, USA
E-mail: alexei@lanl.gov

References

1. V. I. Krinsky, Ed., **Self-Organization: Autowaves and Structures far from Equilibrium,** New York, Springer-Verlag, 1984.
2. L. O. Chua, **CNN: A Paradigm for Complexity,** World Scientific, 1998.
3. A. N. Skourikhine, *Parallel image processing with autowaves: segmentation and edge extraction,* **Proc. of the 4th World Multiconference on Systemics, Cybernetics and Informatics 5,** pp. 302-305, 2000.
4. A. N. Skourikhine, L. Prasad, and B. Schlei, *Neural network for image segmentation,* **Proc. SPIE 4120,** pp. 28-35, 2000.
5. V. G. Yakhno, *Dynamics of autowave processes in neuron-like systems and CNN technology,* **Proc. of the IEEE Int'l Workshop on Cellular Neural Networks and Their Applications,** pp. 33-38, 2000.
6. T. Lindblad and J. Kinser, **Image Processing using Pulse-Coupled Neural Networks,** Springer-Verlag, 1998.

Change detection in aerial stereo pairs at different dates

continued from p. 2

operator will need in order to make an efficient update (focusing on regions with a single change, focusing on areas with many changes, etc.).

In order to improve the updating process for high accuracy databases, we will have to improve our classification step: matching small graphs will not be sufficient in order to take a decision. We will have to compare internal structures of the buildings. This comparison will be accomplished by inter-date 3D feature comparison (segments, planar patches, etc.).

Franck Jung
Institut Geographique National
Laboratoire MATIS
2/4, avenue Pasteur
94165 Saint-Mande cedex, France
Phone: +33 1 43 98 84 36
Fax: +33 1 43 98 85 81
E-mail: franck.jung@ign.fr

References

1. J. Jensen, *Principles of change detection using digital remote sensor data,* in **Integration of Geographic Information System and Remote Sensing,** Star, Estes, McGwire (eds.), Cambridge Univ. Press, 1997.
2. C. Baillard and H. Maitre, *3D Reconstruction of urban scenes from aerial stereo imagery: a focusing strategy,* **Computer Vision and Image Understanding 13,** pp. 244-258, 1999.
3. F. Jung, *Detecting new buildings from aerial stereo pairs at different dates,* **Proc. SPIE 4472,** 2001.
4. N. Paparoditis, M. Cord, M. Jordan, and J.-P. Cocquerez, *Building detection and reconstruction from mid and high-resolution aerial imagery,* **Computer Vision and Image Understanding 72** (2), 1998.

Image restoration: beyond wavelets

continued from p. 3

form will be reproduced by applying the same transform to the solution. Using this approach, the residual is much better, and features cannot be detected by eye any longer. This is not the case for either the wavelet or the curvelet filtering. The combined filtering leads to a real improvement both in terms of signal-to-noise ratio and visual appearance. Furthermore, it arguably challenges the eye in being able to distinguish structure/features from residual images from real image data (at least for the range of noise levels that was considered here). Single transforms cannot manage this.[7]

The complementarity of the different transforms can also be used in order to separate the different components contained in an image. In Reference 8 we proposed the Combined Transform Method, which allows us to represent on different bases simultaneously.

Figure 3 illustrates the result in the case where the input image contains only lines and Gaussians. Two transform operators were used, the wavelet transform and the ridgelet transform. The first is well adapted to the detection of Gaussian due to the isotropy of the wavelet function,[1] while the second is optimal to represent lines.[5] Figure 4 shows the original image, and reconstructed image from the wavelet coefficients, and the reconstructed image from the ridgelet coefficients. The addition of both reconstructed images reproduced the original.[9]

Figure 3. Left: original image containing lines and Gaussians. Middle: reconstructed image for the wavelet coefficient. Right: reconstructed image from the ridgelet coefficients.

Jean-Luc Starck
DAPNIA/SEI-SAP
CEA-Saclay
91191 Gif sur Yvette, France
E-mail: starck@discovery.saclay.cea.fr

References

1. J. Starck, F. Murtagh, and A. Bijaoui, **Image Processing and Data Analysis: The Multiscale Approach,** Cambridge University Press, Cambridge, UK, 1998.
2. **MR/1. Multiresolution Image and Data Analysis Software Package, Version 2.0,** Multi Resolutions Ltd., 2001. http://www.multiresolution.com
3. **Wavelab 802 for Matlab5.x,** 2001. http://www-stat.stanford.edu/~wavelab/.
4. E. J. Candès and D. L. Donoho, *Curvelets—a surprisingly effective nonadaptive representation for objects with edges*, **Curve and Surface Fitting: Saint-Malo 1999,** A. Cohen, C. Rabut, and L. Schumaker, eds., Vanderbilt University Press, Nashville, TN, 1999.
5. E. Candès and D. L. Donoho, *Ridgelets: the key to high dimensional intermittency?*, **Phil. Trans R. Soc. Lond. A 357,** pp. 2495-2509, 1999.
6. J. L. Starck, E. Candès, and D. L. Donoho, *The curvelet transform for image denoising*, **IEEE Trans. on Image Processing,** 2001, *to appear.*
7. J. L. Starck, D. L.Donoho, and E. Candès, *Very high quality image restoration*, **Proc. SPIE 4478,** 2001.
8. J. L. Starck, *Detection of anisotropic features by the curvelet transform*, **Proc. SPIE 4477,** 2001.
9. http://www-stat.stanford.edu/~jstarck

New distortion measures for compression of remote-sensing data

continued from p. 7

tics of the signal can lead to TDOA accuracies that are up to three times better than is possible with minimum MSE approaches.

Filtering/decimation is used to compress the signal by reducing the sampling rate used at the expense of reducing the signal's bandwidth and hence degrading the TDOA accuracy: quantization is used to compress by reducing the number of bits per sample at the expense of decreasing the signal-to-noise ratio (SNR) and hence degrading the TDOA accuracy. Our method determines how to optimally balance the application of filtering/decimation and quantization to achieve optimize TDOA accuracy for a given data rate. Curves of TDOA accuracy versus rate are given in Figure 1.

We have also developed approaches more advanced than simple filtering and decimation. The general goal is the following, expressed here as transform coding with a non-MSE distortion. Given some signal decomposition of the signal to be compressed, we wish to select which coefficients should be coded and transmitted to achieve a desired rate-distortion goal where distortion is a non-MSE measure that captures the structure inherent in the CRB. In general, this selection process is quite difficult because of the nonlinear, nonmonotonic relationship between the coefficients and the RMS bandwidth, and the fact that removing a coefficient affects both the RMS bandwidth and the SNR. We are currently investigating the use of genetic algorithms to accomplish this optimal choice when using a wavelet transform as the signal decomposition. For now, though, a sub-optimal method has been developed that more heavily weights the bandwidth-important high frequencies. Results are given in Figure 2.

Mark L. Fowler
State University of New York at Binghamton
Department of Electrical and Computer Engineering
P.O. Box 6000
Binghamton, NY 13902
E-mail: mfowler@binghamton.edu

References

1. A. Zhang and T. Berger, *Estimation via compressed information*, **IEEE Trans. Info. Theory 34** (2), pp. 198211, March 1988.
2. Mark L. Fowler, *Coarse quantization for data compression in coherent location systems*, **IEEE Trans. Aerosp. and Elect. Systems 36** (4), pp. 1269–1278, October 2000.
3. Mark L. Fowler, *Decimation vs. quantization for data compression in TDOA systems*, **Proc. SPIE 4122**, pp. 56–67, 2000.
4. Mark L. Fowler, *Non-MSE wavelet-based data compression for emitter location*, **Proc. SPIE 4475,** 2001.

Road sign detection for intelligent transportation systems

continued from p. 12

bank of nonlinear filters. Images of a true target captured from different distances constitute the set of filters in the bank. Apart from locating a true sign, this method allows us to approximately determine the distance between it and the acquisition system.[1]

In order to detect slightly-tilted road signs, certain tolerance to in-plane rotations has to be considered. In-plane rotation invariance is achieved by rotating the input scene. Recognition results obtained by this method are compared to results obtained for composite nonlinear filters.[2] Composite filters are constructed by using digital, rotated versions of the reference target. In-plane rotation of the input scene allows better detection results than composite filters. Moreover, in the design of composite filters, the maximum number of images included in a composite filter is limited. The range of the input scene rotation, on the other hand, can be determined based on the application. Using composite nonlinear filters rather than using individual filters in the filter bank can satisfy tolerance requirements for out-of-plane rotation of the targets.[2] By using the same procedure, tolerance to some information included in a given road sign (for instance, the number of a speed-limit sign) is also achieved.[2] In both cases, Fourier-plane nonlinear filters[4,5] are used as composite filters. Finally, as a consequence of using a nonlinear processor, the recognition system has certain tolerance to illumination fluctuations.[1,3]

The entire recognition system has been tested with real still images, as well as with video sequences provided by Connecticut Department of Transportation. Scenes were captured in real environments, with cluttered backgrounds, and contained many distortions simultaneously. Figure 1(a) corresponds to an analyzed scene that includes two stop signs to be detected. These signs are located on both sides of the road, have different illumination and are partially in-plane and out-of-plane rotated. Moreover, the stop sign on the right has a non-uniform illumination, due to shadows caused by the leaves, and has been vandalized. The background of the picture is quite cluttered and there are areas with larger energy than the energy of stop signs.

Two high intensity correlation peaks appear in the output plane, see Figure 1(b), and they coincide with the position of the two true targets in the scene. The image included in Figure 2(a) corresponds to an example of recognition of a speed-limit sign, along with the rejection of objects with similar energy. A high and sharp peak allows location of the target, whereas no false alarms appear: see Figure 2(b). Other distorted speed limits with different speed-limit numbers have also been detected by using the same processor and filter bank.[2] Finally, we demonstrated the rejection of a false sign. The analyzed image shown in Figure 3(a) contains a false target that is perfectly discriminated from the speed-limit target, even though it has a high similarity in shape with respect to the object to be recognized: see Figure 3(b).

[†]Elisabet Perez, [‡]Maria-Albertina Castro and [‡]Bahram Javidi
[†]Optics and Optometry Department
Polytechnic University of Catalunya
Violinista Vellsola, 37
08222 Terrassa, Spain
Tel: +34 93 739 8339
Fax: +34 93 739 8301
E-mail: eperez@oo.upc.es

[‡]Electrical and Computer Engineering Department
University of Connecticut
260 Glenbrook Road
U-157, Storrs
CT 06269-2157, USA
Tel: +1 860 486 2867
Fax: +1 860 486 2447
E-mail: bahram@engr.uconn.edu

References

1. E. Pérez and B. Javidi, *Image processing for intelligent transportation systems: application to road sign recognition,* **Smart Imaging Systems (PM91),** B. Javidi Ed., p. 207, SPIE Press, Bellingham, Washington, 2001.
2. E. Pérez, M.-A. Castro, and B. Javidi, *Strategies for detection of distorted road signs in background noise,* **Proc. SPIE 4471,** 2001 (in press).
3. B. Javidi, *Nonlinear joint power spectrum based optical correlation,* **Appl. Opt. 28,** p. 2358, 1989.
4. B. Javidi and D. Painchaud, *Distortion-invariant pattern recognition with Fourier-plane nonlinear filters,* **Appl. Opt. 35,** p. 318, 1996.
5. B. Javidi, W. Wang, and G. Zhang, *Composite Fourier-plane nonlinear filter for distortion-invariant pattern recognition,* **Opt. Eng. 36,** p. 2690, 1997.

CALENDAR

2002

Photonics West 2002
19–25 January
San Jose, CA USA
Includes International Symposia on
BiOS 2002: Biomedical Optics and Applications
Electronic Imaging 2002: Science and Technology
LASE 2002: Lasers and Applications
Optoelectronics 2002
Program • Exhibition
http://spie.org/conferences/programs/02/pw/

Medical Imaging
23–28 February
Diego, California USA
Call for Papers • Exhibition
http://spie.org/conferences/programs/02/mi/

AeroSense 2002
1–5 April
Orlando, FL USA
Call for Papers
http://spie.org/conferences/programs/02/or/

International Symposium on Optical Science and Technology
SPIE's 47th Annual Meeting
7–11 July
Seattle, WA USA
Call for Papers • Exhibition
http://spie.org/conferences/calls/02/am/

25th International Congress on High Speed Photography and Photonics
29 September–3 October
Beaune, France
SPIE will publish Proceedings
Call for Papers
http://www-le2i.u-bourgogne.fr/hspp2002/

Asian Symposium on Biomedical Optics and Photomedicine
21–23 October
Sapporo, Japan
SPIE is a cooperating organization
Program
http://imd-www.es.hokudai.ac.jp/bopm2002/

For More Information Contact
SPIE • PO Box 10, Bellingham, WA 98227-0010
Tel: +1 360 676 3290 • Fax: +1 360 647 1445 • E-mail: spie@spie.org
Web: www.spie.org

Join the SPIE/IS&T Technical Group

...and receive this newsletter

This newsletter is produced twice yearly and is available only as a benefit of membership in the SPIE/IS&T Electronic Imaging Technical Group.

IS&T—The Society for Imaging Science and Technology has joined with SPIE to form a technical group structure that provides a worldwide communication network and is advantageous to the memberships.

Join the Electronic Imaging Technical Group for US$30. Technical Group members receive these benefits:

- *Electronic Imaging* Newsletter
- SPIE's monthly publication, *oe magazine*
- annual list of Electronic Imaging Technical Group members
- discounts on registration fees for IS&T and SPIE meetings and on books and other selected publications related to electronic imaging.

People who are already members of IS&T or SPIE are invited to join the Electronic Imaging Technical Group for the reduced member fee of US$15.

Please Print ☐ Prof. ☐ Dr. ☐ Mr. ☐ Miss ☐ Mrs. ☐ Ms.

First (Given) Name _____ Middle Initial _____

Last (Family) Name _____

Position _____

Business Affiliation _____

Dept./Bldg./Mail Stop/etc. _____

Street Address or P.O. Box _____

City _____ State or Province _____

Zip or Postal Code _____ Country _____

Phone _____ Fax _____

E-mail _____

Technical Group Membership fee is $30/year, or $15/year for full SPIE and IS&T Members.

Amount enclosed for Technical Group membership **$** _____

❑ **I also want to subscribe to IS&T/SPIE's *Journal of Electronic Imaging* (JEI)** **$** _____
(see prices below)

Total $ _____

❑ **Check enclosed.** Payment in U.S. dollars (by draft on a U.S. bank, or international money order) is required. Do not send currency. Transfers from banks must include a copy of the transfer order.

❑ **Charge to my:** ❑ **VISA** ❑ **MasterCard** ❑ **American Express** ❑ **Diners Club** ❑ **Discover**

Account # _____ Expiration date _____

Signature _____
(required for credit card orders)

Reference Code: 2819

JEI 2002 subscription rates (4 issues):	**U.S.**	**Non-U.S.**
Individual SPIE or IS&T member	$ 55	$ 55
Individual nonmember and institutions	$245	$265

Your subscription begins with the first issue of the year. Subscriptions are entered on a calendar-year basis. Orders received after 1 September 2002 will begin January 2003 unless a 2002 subscription is specified.

Send this form (or photocopy) to:
SPIE • P.O. Box 10
Bellingham, WA 98227-0010 USA
Tel: +1 360 676 3290
Fax: +1 360 647 1445
E-mail: membership@spie.org

Please send me:
☐ **Information about full SPIE membership**
☐ **Information about full IS&T membership**
☐ **Information about other SPIE technical groups**
☐ **FREE technical publications catalog**

Electronic Imaging

The *Electronic Imaging* newsletter is published by SPIE—The International Society for Optical Engineering and IS&T—The Society for Imaging Science and Technology. The newsletter is the official publication of the International Technical Group on Electronic Imaging.

Editor / Technical Group Chair	Arthur Weeks
Technical Editor	Sunny Bains
Managing Editor / Graphics	Linda DeLano
Advertising Sales	Roy Overstreet

Articles in this newsletter do not necessarily constitute endorsement or the opinions of the editors, SPIE, or IS&T. Advertising and copy are subject to acceptance by the editors.

SPIE is an international technical society dedicated to advancing engineering, scientific, and commercial applications of optical, photonic, imaging, electronic, and optoelectronic technologies.

IS&T is an international nonprofit society whose goal is to keep members aware of the latest scientific and technological developments in the fields of imaging through conferences, journals and other publications.

SPIE—The International Society for Optical Engineering, P.O. Box 10, Bellingham, WA 98227-0010 USA. Tel: +1 360 676 3290. Fax: +1 360 647 1445. E-mail: spie@spie.org.

IS&T—The Society for Imaging Science and Technology, 7003 Kilworth Lane, Springfield, VA 22151 USA. Tel:+1 703 642 9090. Fax: +1 703 642 9094.

Want it published?

Submit work, information, or announcements for publication in future issues of this newsletter by e-mail to Sunny Bains. For full submission information, length and format required, deadline, etc., please go to:

http://www.sunnybains.com/newslet.html.

All materials are subject to approval and may be edited. Please include the name and e-mail address of someone who can answer technical questions. Calendar listings should be sent at least eight months prior to the event.

EIONLINE

New Electronic Imaging Web Discussion Forum launched

You are invited to participate in SPIE's new online discussion forum on Electronic Imaging. The **INFO-EI** mailing list is being "retired" as we move the discussions to the more full-featured web forums. We hope you will participate. To post a message, log in to create a user account. For options see "`subscribe to this forum.`"

You'll find our forums well-designed and easy to use, with many helpful features such as automated email notifications, easy-to-follow "threads," and searchability. There is a full FAQ for more details on how to use the forums.

Main link to the new Electronic Imaging forum: `http://spie.org/app/forums/tech/`

Related questions or suggestions can be sent to `forums@spie.org`.

Road sign detection for intelligent transportation systems

Development of safety systems is in increasing demand. A safety system based on a pattern recognition processor could be installed in vehicles in order to automatically detect and identify road signs. Afterwards, the recognition system could make an objective decision according to the information detected. One of the greatest difficulties on achieving this goal lies in the number of different distortions that may simultaneously modify the reference sign. Variations in scale, in-plane and out-of-plane rotations, background clutter, partially occluded signs, and variable illumination, are some examples of distortions that can affect road signs.

Different approaches to obtain distortion-tolerant systems have been developed in the field of pattern recognition. In general, a given recognition technique is designed to provide satisfactory results when dealing with a particular distortion of the object. However, the same strategy usually gives poorer results if another type of distortion has affected the object. We have carried out analysis and comparison of different techniques. By combining various strategies, we obtained a recognition system that is simultaneously scale-invariant and tolerant to slight tilts or out-of-plane rotations (due to different view angles of the acquisition system). Tolerance to illumination fluctuations is also needed to enable a recognition system to work under different illumination or weather conditions. Finally, robustness to a cluttered background is important for a road sign recognition processor that analyzes images captured in real environments.

Recently, we proposed a road sign recognition system[1,2] based on a nonlinear processor.[3] The processor performs several nonlinear correlations between different input scenes and a set of reference targets. Multiple correlation results are then processed to give a single recognition output. A learning algorithm is used to establish a threshold value that determines whether or not any object contained in an input scene is similar to the target.

To achieve detection of road signs even when the acquisition system is in motion, scale-invariance is provided to the processor by means of a

continued on p. 10

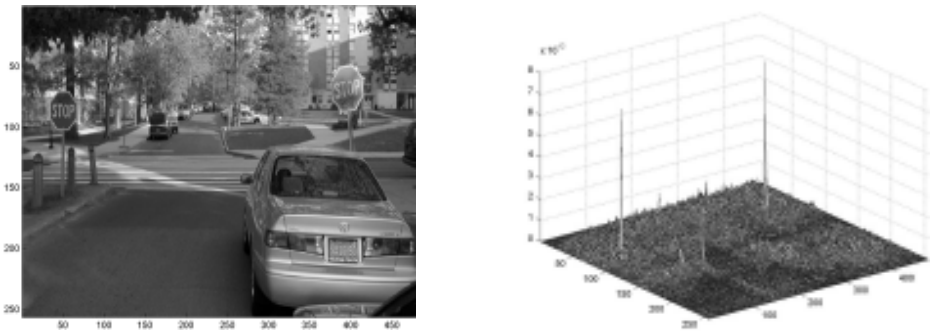


Figure 1. Recognition of stop signs using the described distorted-tolerant system. (a) Input scene. (b) 3D representation of the output plane.

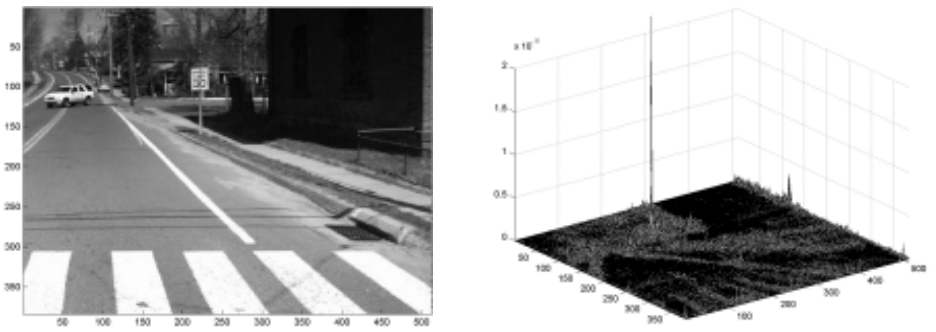


Figure 2. Recognition of a 30mph speed-limit sign, on a cluttered background, using the described distorted-tolerant system. (a) Input scene. (b) 3D representation of the output plane.

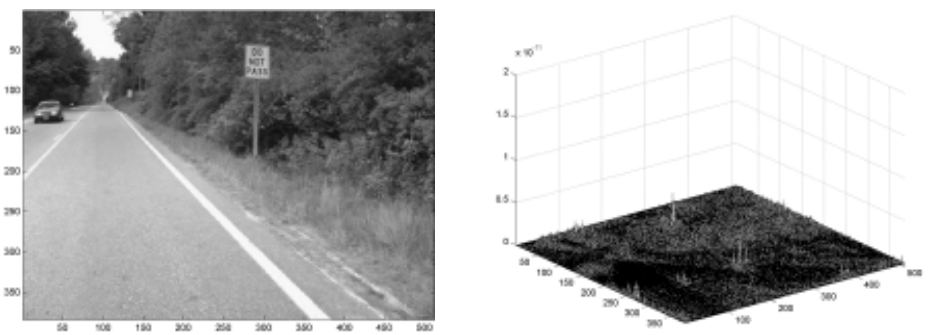


Figure 3. Recognition results for the described distorted-tolerant system when a false sign (do not enter sign) is analyzed. a) Input scene. b) 3D representation of the output plane.

SPIE Society of Photo-Optical Instrumentation Engineers

P.O. Box 10 • Bellingham, WA 98227-0010 USA

Change Service Requested

DATED MATERIAL

Non-Profit Org.
U.S. Postage Paid
Society of
Photo-Optical
Instrumentation
Engineers