POINTWISE SHAPE-ADAPTIVE DCT DENOISING WITH STRUCTURE PRESERVATION IN LUMINANCE-CHROMINANCE SPACE

Alessandro Foi, Vladimir Katkovnik, Karen Egiazarian

Institute of Signal Processing, Tampere University of Technology, Tampere, FINLAND
firstname.lastname@tut.fi

ABSTRACT

The shape-adaptive DCT (SA-DCT) [12, 13] is a low-complexity transform which can be computed on a support of arbitrary shape. Particularly suited for coding image patches in the vicinity of edges, the SA-DCT has been included in the MPEG-4 standard [7].

The use of this shape-adaptive transform for grayscale image denoising has been recently proposed [3, 2], showing a remarkable performance.

In this paper we extend this approach to color filtering in luminance-chrominance space, exploiting the structural information obtained from the luminance channel to drive the shape-adaptation for the chrominance channels.

Several simulation experiments attest the advanced performance of the proposed color denoising method. The visual quality of the estimates is high, with sharp detail preservation, clean edges, and without unpleasant ringing artifacts introduced by the fitted transform.

Besides noise removal, the proposed method is also effective in dealing with those artifacts which are often encountered in compressed images and videos. Blocking artifacts are suppressed while salient image features are preserved. The SA-DCT filtering used for the chrominance channels allows to faithfully reconstruct the missing structural information of the chrominances, thus correcting color-bleeding artifacts. Being based on the SA-DCT (which is implemented as standard in modern MPEG hardware), the proposed method can be integrated within existing video platforms as a pre- or post-processing filter.

1. INTRODUCTION

The shape-adaptive DCT (SA-DCT) [12, 13] is a generalization of the usual separable block-DCT (B-DCT). While retaining the same computational complexity of the separable B-DCT, the SA-DCT can be computed on a support of arbitrary shape. The SA-DCT has been originally developed for coding non-rectangular image patches near the border of image objects, in such a way to minimize the stored information and to avoid the ringing artifacts (Gibbs phenomenon) that would appear in correspondence with strong edges. Because of its low-complexity, the near-optimal decorrelation and energy compaction properties (e.g. [12],[5]), and its backward-compatibility with the B-DCT, the SA-DCT has been included in the MPEG-4 standard [7], where it is used for the coding of image segments that lie on the video-object's boundary.

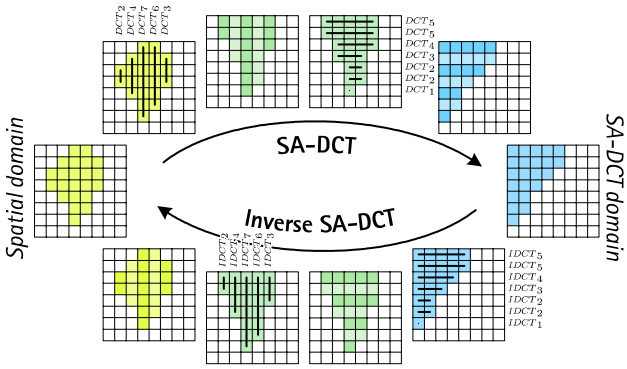

Fig. 1. Illustration of the Shape-Adaptive DCT transform and its inverse. Transformation is computed by cascaded application of one-dimensional varying-length DCT transforms, along the columns and along the rows.

The SA-DCT is computed by cascaded application of one-dimensional varying-length DCT transforms first on the columns and then on the rows that constitute the considered region, as shown in Figure 1.

The recent availability of low-power hardware SA-DCT platforms (e.g. [1],[6]) makes this transform an appealing choice for many image- and video-processing tasks.

The first attempt to use of SA-DCT for image denoising was reported by the authors in [3]. The original version of the method has been refined and a shape-adaptive transform-domain deconvolution filter for image deblurring has also been developed [2]. These denoising and deblurring algorithms demonstrate a remarkable performance, usually outperforming the best methods known to the authors.

In this paper we extend the SA-DCT filtering approach to color denoising. The proposed method is based on a luminance-chrominance color-transformation and exploits the structural information obtained from the luminance chan-

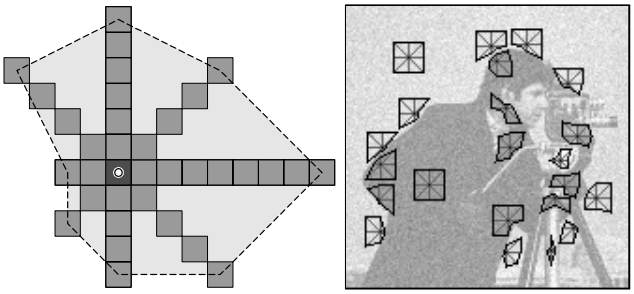

Fig. 2. The anisotropic neighborhood is constructed as the polygonal hull of the adaptive-scale directional kernels supports (left). Some examples of the anisotropic neighborhoods for the noisy *Cameraman* image (right).

nel to drive the shape-adaptation for the chrominance channels. Such adaptation strategy enables accurate preservation and reconstruction of image details and structures and yields estimates with a very good visual quality.

Since the SA-DCT is implemented as standard in modern MPEG hardware, the proposed method can be integrated within existing video platforms as a pre- or post-processing filter.

The paper is organized as follows. In the next section we give an overview on the basic anisotropic LPA-ICI + SA-DCT method [3, 2]. In Section 3 we present its extension to color denoising, describing the employed color-space transformations and discussing the structural constraints which are imposed on the chrominances. Two sections are devoted to experimental results: several examples of additive Gaussian white noise removal are shown in Section 4; the application of the proposed method to deblocking and deringing from block-DCT compressed data is presented in the Section 5.

2. ANISOTROPIC LPA-ICI-DRIVEN SA-DCT DENOISING

In this section we give a brief overview on the original method for (grayscale) image denoising. We refer the reader to [2] (and references therein) for a more formal, complete, and rigorous description.

In our approach we use the SA-DCT in conjunction with the anisotropic LPA-ICI technique [4]. By comparing varying-scale directional kernel estimates, the technique adaptively selects, for each point in the image, a set of directional adaptive-scales. The length of the support of the corresponding adaptive-scale kernels define the shape of the transform's support in a pointwise-adaptive manner. Examples of such neighborhoods are shown in Figure 2.

For each one of these neighborhoods a SA-DCT is performed. The hard-thresholded SA-DCT coefficients are used to reconstruct a local estimate of the signal within the adaptive-shape support. By using the adaptive neighborhoods as support for the SA-DCT, we ensure that data are represented sparsely in the transform domain, allowing to effectively separate signal from noise using hard-thresholding.

Since supports corresponding to different points are in general overlapping (and thus generate an overcomplete representation of the signal), the local estimates are averaged together using adaptive weights that depend on the local estimates' statistics. In this way we obtain and adaptive estimate of the whole image.

Once this global estimate is produced, it is used as reference estimate for an empirical Wiener filter in SA-DCT domain. Following the same adaptive averaging procedure as for hard-thresholding, we arrive to the final Anisotropic LPA-ICI-driven SA-DCT estimate.

We remark that the LPA-ICI technique is fast, since it is based on convolutions against one-dimensional kernels for a very limited number of directions. It constitutes a negligible overhead for the whole SA-DCT filtering algorithm.

3. POINTWISE SA-DCT DENOISING IN LUMINANCE-CHROMINANCE SPACE

Let $y = [y_R, y_G, y_B]$ be the original color image in the *RGB* color space. We consider noisy observations $z = [z_R, z_G, z_B]$ of the form $z_C = y_C + n_C$, $C = R, G, B$, where the noise $n = [n_R, n_G, n_B]$ is independent Gaussian, $n_C(\cdot) \sim \mathcal{N}(0, \sigma_C^2)$.

In order to deal with color images, we first perform a color-space transformation, aiming at reducing the strong correlation between channels which is typical of the *RGB* space. In particular, we consider the "*opponent*" and the *YUV* color spaces [9]. Up to some normalization, the transformation to these color spaces can be expressed by multiplication of a column vector with the R, G, and B components against one of the matrices

$$\mathbf{A}_{opp} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} & \frac{-\sqrt{2}}{3} & \frac{1}{3\sqrt{2}} \end{bmatrix}, \quad \mathbf{A}_{yuv} = \begin{bmatrix} 0.30 & 0.59 & 0.11 \\ -0.18 & -0.36 & 0.54 \\ 0.51 & -0.43 & -0.08 \end{bmatrix}.$$

Although purists may consider it an abuse of terminology, we call "luminance" and "chrominances" not only the components the *YUV* space, but also those of the opponent color space. We denote the luminance channel as Y, and the chrominances as U and V.

In such luminance-chrominance decompositions, the original inter-channel correlation of the *RGB* space is captured into the luminance channel – which thus enjoys a better signal-to-noise ratio (SNR) – whereas the chrominance channels contain the differential information among the *RGB* channels. It is interesting to observe that the rows of the matrix $\mathbf{A}_{opp}$ actually correspond to a mean filter and to a first- and second-order finite-difference filters.

The noise variances of the Y, U, and V channels can be calculated as the elements of the vector $[\sigma^2_Y\ \sigma^2_U\ \sigma^2_V] = [\sigma^2_R\ \sigma^2_G\ \sigma^2_B]\mathbf{A}^{T2}$, where σ^2_R, σ^2_G, and σ^2_B are the noise variances for the R, G, and B channels and $\mathbf{A}^{T2}$ is the transposed color transformation matrix with all elements squared.

Ideally, the Y, U, and V channels are considered as independent, with most of the valuable information typically collected into the luminance channel Y. Therefore, the common approach for color denoising in luminance-chrominance space is to filter the three channels separately and independently one from the other.

However, when considering natural images, the different color channels always share some common features which are inherited from the structures and from the objects depicted in the original image. In particular, it can be observed that along the objects' boundaries all color channels usually exhibit some simultaneous discontinuities or sharp transitions.

We exploit this kind of structural correlation by imposing that the three transform's supports which are used for the filtering of the Y, U, and V channels at a particular location have the same adaptive shape. In practice, we use for all three channels the adaptive neighborhoods defined by the anisotropic LPA-ICI for the luminance channel.

Such a constraint make so that whenever some structure is detected (by the LPA-ICI), it is taken into account (and thus preserved) for the filtering of all three channels. Restricted to the adaptive supports, however, the channels are assumed as independent, and thus the transform-domain hard-thresholding and Wiener filtering are still performed for each channel independently from the others.

After the filtering of the three channels, inverse color-transformation returns the denoised image in the *RGB* space.

The better SNR of the luminance and its higher "information content" are the two main reasons why it is in this channel that we look for structures. There are also other reasons. In natural images it often happens that uniformly colored objects present luminance variations due to non-uniform illumination or shadowing: such transitions cannot be detected from the chrominances. On the other hand, it is quite rare that abrupt changes appear in the chrominances and not in the luminance. Therefore, it is sufficient to perform the LPA-ICI adaptive-scale selection on the luminance channel only.

As shown in Section 5, the proposed method not only accurately preserves the structures in the different channels at low SNR levels, but is also able to *reconstruct missing structural information* in the chrominance channels.

It is striking that this can be done *without* assuming any form of correlation between the intensities nor between the transform's coefficients corresponding to different channels (because the outcome of hard-thresholding and Wiener-filtering depends solely on the data in the channel being filtered). Roughly speaking, the overall approach can be interpreted as channel-by-channel independent filtering in a structure-optimized overcomplete transform-domain. Such optimization of the transform is done with respect to the structures detected in the luminance channel.

Fig. 3. Fragments of the original *F-16* image (top-left), of its noisy observation (σ=30, PSNR=18.59) (top-right) and of two denoised estimates: our SA-DCT estimate (bottom-left), and the ProbShrink-MB [8] estimate (bottom-right). The PSNR for the two estimates is 31.59dB and 30.50dB, respectively.

4. DENOISING EXPERIMENTS

In order to assess the denoising perfomance of the proposed method, we performed a number of denoising experiments over standard color test-images. For these experiments the variance of the additive Gaussian noise is set to be the same for all *RGB* color channels, $\sigma^2_R = \sigma^2_G = \sigma^2_B = \sigma^2$. Filtering is performed after transformation to the opponent color space. Table 1 gives the PSNR results for the denoising of the *Lena*, *Peppers*, *Baboon*, *House* and *F-16* color test-images over a wide range of values of σ.

In Table 2 we compare our results against those by other state-of-the-art methods, as reported in [8]. In particular, the vector-based minimum-mean-squared-error estimator (VMMSE) [11], the multiband wavelet thresholding (MBT) [10], and the ProbShrink-multiband wavelet algorithm [8] are considered for comparison.

Let us note that the reference methods which are in-

Fig. 4. Fragments from the noisy (σ=25, PSNR=20.18dB) and denoised *Peppers* image (PSNR=30.90dB), obtained using the proposed SA-DCT algorithm.

Method σ	*Lena* 512x512 RGB				*Peppers* 512x512 RGB				*Baboon* 512x512 RGB			
	10	15	20	25	10	15	20	25	10	15	20	25
Anis. LPA-ICI + SA-DCT	**34.95**	**33.58**	**32.61**	**31.85**	**33.70**	**32.42**	**31.57**	**30.90**	30.62	**28.33**	**26.89**	**25.86**
ProbShrink-MB [8]	34.60	33.03	31.92	31.04	33.44	32.05	31.12	30.35	30.17	27.83	26.38	25.27
VMMSE [11]	34.02	31.89	30.24	28.88	33.12	31.13	29.67	28.45	**30.68**	28.24	26.63	25.36
MBT [10]	33.84	32.29	31.14	30.15	31.19	30.22	29.45	28.77	28.50	26.78	25.53	24.56

Table 2. PSNR (dB) comparison table for the denoising of the *Lena*, *Peppers*, and *Baboon* color test images with different levels of Gaussian noise.

σ	*Lena*	*Peppers*	*Baboon*	*House*	*F-16*
10	34.95	33.70	30.62	35.67	36.41
15	33.58	32.42	28.33	34.09	34.67
20	32.61	31.57	26.89	32.97	33.41
25	31.85	30.90	25.86	32.12	32.42
30	31.21	30.33	25.07	31.39	31.59
35	30.65	29.81	24.44	30.74	30.88
50	29.27	28.53	23.03	29.13	29.19
75	27.77	27.07	21.46	27.39	27.43

Table 1. Color image denoising performance as PSNR (dB) for the proposed SA-DCT algorithm.

cluded in the Table 2 are all multiband or vector methods, which are specifically designed for the denoising of color or multispectral images. Such algorithms filter simultaneously all channels, exploiting the possible inter-channel correlation, and are thus inherently superior to the simpler strategy where a scalar (grayscale) denoising filter is used independently for each separate channel.

We remark that although in our approach the adaptive supports for the SA-DCT at particular location are the same for all three channels, the SA-DCT-domain filtering is performed for each channel independently. Nevertheless, our results are very competitive and the comparison table shows that in fact the proposed technique outperforms the other reference methods. Additional improvement can be expected if the current channel-by-channel filtering is replaced by a more sophisticated shape-adaptive transform which filters – in vectorial form – all channels simultaneously. It must be noted, however, that such a transform would be more complex and also drastically different from the scalar SA-DCT which is considered within the MPEG-4 standard.

Similarly to the case of grayscale denoising, reported in [2], the denoised color estimates produced by our adaptive algorithm are visually very good. A close inspection to Figure 3 and 4 may reveal the outstanding preservation of sharp details achieved by the shape-adaptive transform. At the same time, almost no visible artifacts are present. Other transform-based methods often introduce noticeable blurriness and overshooting on the edges and unpleasant spurious oscillations in smooth areas. These artifacts do not appear in our estimates thanks to the pointwise-adaptive selection of the transform's support.

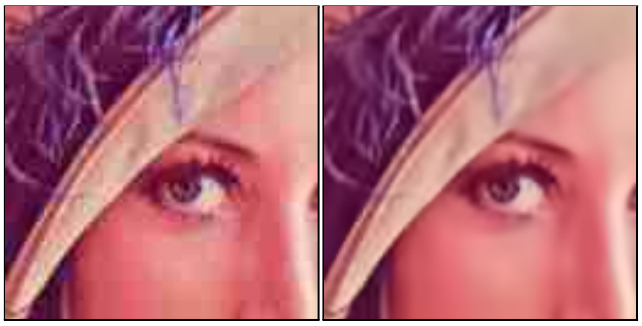

Fig. 5. Fragment of the JPEG-compressed *Lena* image (quality=20, 0.38bpp, PSNR=29.83dB) and of its SA-DCT filtered estimate (PSNR=31.00dB).

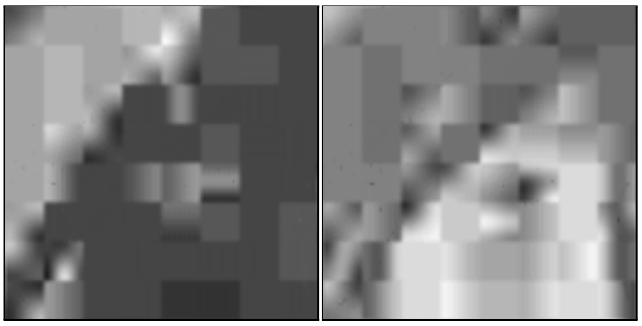

Fig. 6. Detail of the U and V chrominances of the JPEG-compressed *Lena* image shown in Figure 5(left). Only very few DCT harmonics survived the aggressive quantization, and the structural information is almost completely lost.

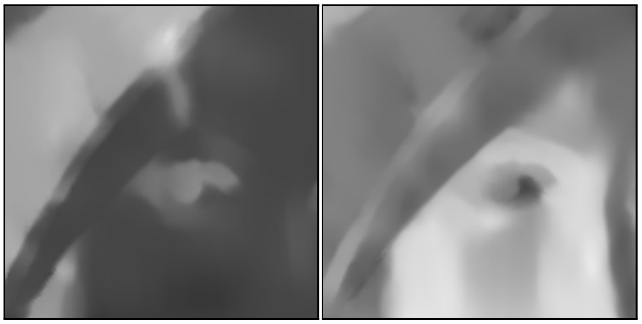

Fig. 7. The chrominances from Figure 5 after reconstruction by SA-DCT filtering. The blockiness is removed and the structures are faithfully restored.

5. DERINGING AND DEBLOCKING FROM BLOCK-DCT COMPRESSION

We conclude the paper presenting an application of the proposed SA-DCT denoising method to postfiltering of compressed data. As an illustrative example, we consider JPEG-compression, which is based on *YUV* color-transformation and on quantization in B-DCT domain. Nevertheless, our approach is general and can be utilized for post-filtering of other image- or video-compression standards.

While more sophisticated models of B-DCT-domain quantization noise have been proposed by many authors, we model the degradation due to quantization as some additive Gaussian white noise, estimating the appropriate values for σ_Y^2, σ_U^2, σ_V^2 from the corresponding quantization matrices. The procedure is fully automatic and allows to apply the above SA-DCT denoising algorithm as an effective deblocking and deringing filter for B-DCT coded images and videos. The proposed method is particularly relevant for video post-processing, since it can exploit the SA-DCT hardware of MPEG-4 decoders.

A peculiarity of our approach is the ability to filter the chrominance channels restoring the structural information which went lost because of quantization and coarse sampling. Even at relatively high bit-rates, the compression of the chrominance channels is usually quite aggressive. It results in unpleasant color-bleeding artifacts along edges between differently colored objects. At the same time, on smoother areas the uneven hue due to quantization becomes particularly noticeable.

Figure 5(left) shows a fragment of the JPEG compressed *Lena* image (quality=20, 0.38bpp, PSNR=29.83dB). The corresponding U and V chrominance channels are shown in Figure 6. One can barely recognize the salient features of the image, such as the border of the hat or the contours of the eyes and nose. These structures can be faithfully restored by the use of adaptive-shape supports which are determined from the luminance channel, as shown in Figure 7. It is remarkable that even small details such as the iris can be accurately reconstructed from the coarse available information using adaptive transform's supports. The restored color image (PSNR=31.00dB) is shown in Figure 5(right). The ringing and the blocking artifacts disappeared, whereas no details have been oversmoothed, demonstrating the superior adaptivity of the approach. Moreover, thanks to the accurate reconstruction of the structures in the chrominance channels, our estimate does not exhibit any significant chromatic distortion and has a natural appearence.

Although it is well-established that the human visual system is less sensitive to distortions in the chrominances than to those in the luminance, the importance of restoring the chrominances must not be overlooked. In fact, modern image and video compression standards are all designed to exploit the characteristics of the human visual system, and thus adjust the compression rate for the luminance and chrominance channels in such a way to balance the perceptual impact of the distortions among the three channels. Therefore, when visual quality is of concern, the restoration of the different channels deserves equal attention. The coarser quantization of the chrominances makes their accurate restoration a much more difficult and challenging task.

Furthermore, the proper reconstruction of the structural information in the chrominance channels is crucial when

Fig. 8. Postprocessing example: automatic color balance and bicubic interpolation. A detail of the post-processed *Lena* JPEG from Figure 5 is shown on the left. The image on the right has been is obtained by filtering the JPEG with the proposed SA-DCT denoising before the same post-processing.

further processing is applied to the compressed images. In Figure 8(left) we show a fragment of the compressed *Lena* image which has undergone automatic color balance and interpolation. Such simple operations are typically employed in most digital photography printers. Although the compression level is not dramatic, if no filtering is performed the compression artifacts become obvious: the chrominance macroblocks can be distinguished easily and color-bleeding abounds. Contrary to that, the JPEG filtered by the proposed SA-DCT denoising algorithm can be post-processed with no complications, resulting in a final image of significantly superior quality, as shown in Figure 8(right).

More simulation results and the MATLAB software which implements the presented method are available at `http://www.cs.tut.fi/~foi/SA-DCT/`.

6. REFERENCES

[1] Chen, K.-H., J. Guo, J.-S. Wang, C.-W. Yeh, and J.-W. Chen, "An energy-aware IP core design for the variable-length DCT/IDCT targeting at MPEG4 shape-adaptive transforms", *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15, no. 5, pp. 704-714, May 2005.

[2] Foi, A., K. Dabov, V. Katkovnik, and K. Egiazarian, "Shape-adaptive DCT for denoising and image reconstruction", *Proc. SPIE El. Imaging 2006, Image Process.: Algorithms and Systems V*, 6064A-18, January 2006.

[3] Foi, A., V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT as an overcomplete denoising tool", *Proc. 2005 Int. TICSP Workshop Spectral Meth. Multirate Signal Process., SMMSP 2005*, pp. 164-170, Riga, June 2005.

[4] Katkovnik, V., A. Foi, K. Egiazarian, and J. Astola, "Directional varying scale approximations for anisotropic signal processing", *Proc. XII European Signal Process. Conf., EUSIPCO 2004,* pp. 101-104, Vienna, September 2004.

[5] Kauff, P., and K. Schuur, "Shape-adaptive DCT with block-based DC separation and ΔDC correction", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 237-242, 1998.

[6] Kinane, A., A. Casey. V. Muresan, and N. O'Connor, "FPGA-based conformance testing and system prototyping of an MPEG-4 SA-DCT hardware accelerator", *IEEE 2005 Int. Conf. on Field-Programmable Technology, FPT' 05*, Singapore, December 2005.

[7] Koenen, R., "Overview of the MPEG-4 Standard", ISO/IEC JTC1/SC29/WG11 Doc. N3536, July 2000.

[8] Pižurica, A., W. Philips, and P. Scheunders, "Wavelet domain denoising of single-band and multiband images adapted to the probability of the presence of features of interest", *Proc. of SPIE*, *Wavelets XI*, vol. 5914, September 2005.

[9] Plataniotis, K.N., and A.N. Venetsanopoulos, *Color image processing and applications*, Springer-Verlag New York, Inc., New York, NY, 2000

[10] Scheunders P., "Wavelet thresholding of multivalued images", *IEEE Trans. Image Process.*, vol. 13, pp. 475-483, April 2004.

[11] Scheunders P., and J. Driesen, "Least squares interband denoising of color and multispectral images", *Proc. IEEE Int. Conf. Image Process., ICIP 2004*, Singapore, October 2004.

[12] Sikora, T., "Low complexity shape-adaptive DCT for coding of arbitrarily shaped image segments", *Signal Process.: Image Comm.*, vol. 7, pp. 381-395, 1995.

[13] Sikora, T., and B. Makai, "Shape-adaptive DCT for generic coding of video", *IEEE Trans. Circuits and Systems for Video Techn.*, vol. 5, no. 1, pp. 59-62, 1995.