

# Simultaneous Image Denoising and Compression by Multiscale 2D Tensor Voting\*

Yu-Wing Tai    Wai-Shun Tong    Chi-Keung Tang  
Vision and Graphics Group  
The Hong Kong University of Science and Technology  
{yuwing,cstws,cktang}@cs.ust.hk

## Abstract

*In this paper we propose a method that simultaneously performs image denoising and compression by using multiscale tensor voting. Given a real color image, the pixels are first converted into a set of tokens to be grouped by tensor voting, where optimal scales are automatically selected among others for perceptual grouping and faithful reconstruction. Tensor voting at multiple scales are performed at all input tokens to infer the feature grouping attributes such as region-ness, curve-ness, and junction-ness with their optimal scales. We perform experiments on complex real images to demonstrate the robustness of our method.*

## 1. Introduction

The primary hurdle for many algorithms and applications in computer vision is the presence of image noises and ambiguities. Image noises due to quantization errors, sensor errors, limited dynamic range and lens distortion are inevitably introduced during the image formation process. The presence of image ambiguities caused by low contrast, accidental edge alignment, and the presence of multiple feature interpretations at a pixel render the problem of image denoising very difficult.

The bulk majority of work on image denoising and feature/noise disambiguation is based on multiscale analysis, specifically by the construction of scale space. In [1, 2], Lindeberg presented a framework for multiscale image analysis and automatic selection of scale, which is given by selecting the normalized measures of feature strength in the scale space. The Perona-Malik equation [5] deals with the scale space and performs anisotropic diffusion for edge-preserving smoothing. It is observed in [4] that the proper selection of scale is instrumental in grouping relevant information, suppressing irrelevant noises while preserving the necessary image details. Because in general image features occur at multiple scales, the optimal scale of analysis should be different at each pixel in order to achieve the simulta-

neous grouping of smooth features without over-smoothing fine details and the stopping of any smoothing at discontinuities.

To address the problem of scale selection, instead of constructing a scale space or performing PDE-based diffusion, we propose to use tensor voting at multiple scales, where perceptual grouping and faithful reconstruction can be achieved. Further, our faithful reconstruction allows image denoising and image compression be simultaneously performed. Subject to the sampling theory and the use of smooth tensor voting fields, image points (tokens) are grouped in salient feature types at optimal scales where a minimal set of representative tokens is extracted in each scale for achieving a reconstruction that is faithful to the original input without noise. In other words, our approach denoises an image via robust faithful reconstruction, with the additional advantage that a compact (minimal) representation for the noiseless image is also obtained. This paper presents encouraging results on this new idea on image denoising and compression, while quantitative evaluation and comparison with state-of-the-art techniques will be addressed as future work.

## 2. Simultaneous Image Denoising and Compression

In this section, we present a method where image denoising and compression are simultaneously performed on a real color image. To deal with real color images, we need to encode the image pixels into a set of point tokens. Because colors are available at each pixel, color information is used for scale analysis and grouping. We perform multiscale analysis on the local color distribution at each pixel to estimate the optimal scale for each encoded token, and then group them into region interior and region boundary.

### 2.1 Token grouping

We first convert the image pixels into a set of tokens for tensor voting to operate by the following simple color analysis, which is approximate and does not always give good results. Errors are manifested into missing data or outliers.

\*This research is supported by the Research Grant Council of Hong Kong: HKUST6171/03E.

Multiscale tensor voting is used in this paper which is robust against these noises.

1. A range of voting field size (scale) is defined. For each pixel, the local color distribution is modeled at each discrete scale by a Gaussian distribution  $(\mu, \Sigma)$  where  $\mu$  and  $\Sigma$  are the means and covariance matrix respectively, The Gaussian model is inferred by standard techniques (e.g. expectation maximization).
2. For each scale, we measure the token saliency by its conformity to the estimated color model. If the token does not fit to a Gaussian model at any scales, this token is considered to be noise.
3. Given a salient image token, if  $|\Sigma|$  is large, the token is lying near the region boundary where color change is more evident. If  $|\Sigma|$  is small, it indicates that the token is lying in the interior of a smooth color region where the color transition is smooth and the local region can be represented by a single token.

Hence, we define the optimal scale for each token to be largest neighborhood size (scale) that gives a small  $|\Sigma|$ . If the value of  $|\Sigma|$  is large at the smallest scale, these tokens are labeled as region boundary tokens. Else, the pixel is labeled as a region token.

## 2.2 Representative token extraction

The above token grouping stage based on colors is empirical, and therefore errors are present. Now, we extract the representative tokens at each successive scale.

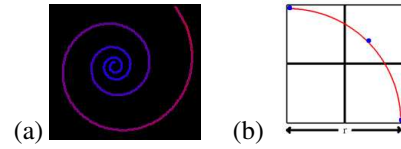
Subject to the Nyquist sampling theory, we define the set of representative tokens to be a minimal set of point tokens that produces a faithful reconstruction by applying a smooth reconstruction filter with the optimal scale at each token. In our case, the tensor voting field is the smooth reconstruction filter, and the optimal scale is defined as the size of the voting field corresponding to the highest normalized feature saliency.

### 2.2.1 Extracting representative curve tokens by curvature grouping

Note that tokens labeled as curve as well as region boundary (where a curve is extracted) are handled in the same way.

Suppose that the underlying finite curve is parameterized by  $\alpha(t), 0 \leq t \leq 1$ . Given a curve token, it is either lying in the curve interior where  $0 < t < 1$ , or a curve endpoint/junction where  $t = 0$  or  $1$ .

The curve tokens corresponding to  $\alpha(0)$  and  $\alpha(1)$  should therefore be extracted as representative tokens because they represent discontinuities. Recall that after tensor voting in the token grouping stage, each token is assigned the endpoint and junction saliencies, which are given by the first order votes received after tensor voting [3]. Hence,



**Figure 1.** (a) Curvature grouping, where the optimal scale is color coded in the spiral shape: red indicates large scale (low curvature) and blue indicates small scale (high curvature). (b) Using the Nyquist sampling theory, the grid size is equal to  $\frac{r}{2}$  where  $r$  is the radius of curvature given by  $\max(C)$  in the curvature group, and  $C$  is the estimated curvature at a curve token.

the extraction is achieved by detecting the maxima in endpoint/junction saliency and performing non-maximal suppression.

For  $\alpha(t), 0 < t < 1$ , the curvature  $\alpha''(t)$  should be considered in order to perform optimal sampling subject to the sampling theory: the larger  $\alpha''(t)$  is, the higher the sampling rate should be and thus more representative tokens should be extracted. Recall that the curvature at a curve token is available after tensor voting at the token grouping stage.

Since the stick voting field prefers smooth and low-curvature connections, tokens with low-curvature feature should be sparsely sampled and vice versa. Therefore, we first group the curve tokens by their curvature values, where each group consists of tokens with similar curvatures defined as  $\max(C)/\min(C) = 2$  and  $C$  is the estimated curvature at each token. Fig. 1(a) shows the result of curvature grouping.

Based on the Nyquist sampling theory for faithful reconstruction and the result of curvature grouping, the optimal size of the reconstruction filter (or voting field) at each token is given by its curvature value. We use a grid-based approach for extracting representative tokens as follows. For each curvature group:

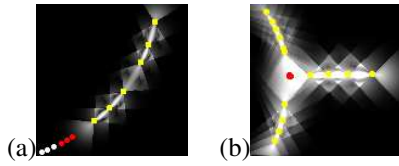
1. We use a 2D grid with the grid size equals to  $\frac{1}{2 \max(C)}$ . Fig. 1(b) shows a typical scenario.
2. For each grid, the token closest to the grid center is extracted as a representative token.

Our grid-based approach is similar in spirit to quadtree decomposition, but is different because our grid size is derived using the curvature at each token, which is related to its optimal scale for faithful reconstruction.

### 2.2.2 Extracting representative region tokens by density grouping

We extract representative region tokens in a way similar to extracting representative curve tokens.

The extraction of representative region tokens is performed at multiple scales. The local density of a region indicates the optimal scale for each region token: the higher



**Figure 5.** Scale transition. (a) Smooth transition. We show the vote saliency received from representative tokens (yellow). In smooth transition, the extrapolated curve is guided by tensor voting field after including tokens at the larger scale (red tokens) and producing smooth connection across a different scale. (b) In discontinuity, curve endpoints and junctions are precisely detected (local maximum in vote saliency) by tensor voting field.

the region density, the smaller the region scale because a smaller voting field is capable of connecting any two region tokens, and hence less representative tokens should be extracted.

Similar to grouping curve tokens into curvature groups by the estimated curvatures, we group region tokens into density groups by their respective local densities. According to the Nyquist rate, the size of the voting field and hence the sampling rate are therefore equal to  $\max(D)/\mathcal{N}$ , where  $\max(D)$  is the maximum density in each density group, and  $\mathcal{N}$  is the neighborhood size for local density estimation. The representative region tokens are similarly extracted using the grid-based approach.

Fig. 2(a) shows the set of representative tokens thus obtained, which is minimal subject to the sampling theory and the smooth voting fields. Fig. 2(b) shows the optimal scales in color codes.

## 2.3 Multiscale feature extraction

We perform multiscale feature extraction and integration. Recall in the previous section:

1. For each curvature group, the optimal scale or size of the voting field is  $\frac{1}{\max(C)}$ .
2. For each density group, the optimal scale or size of the voting field is  $\frac{\max(D)}{\mathcal{N}}$ .

Using the tensor voting fields of the optimal size derived in each feature group, we perform multiscale feature extraction that is faithful to the presumably noiseless input and preserves the necessary details subject to the sampling theory.

### 2.3.1 Curve/region boundary extraction

Because we use the modified marching cubes algorithm [3] to perform curve extraction, given the representative curve tokens in each curvature group, tensor voting with the optimal field size is used to extract the curve(s) in each group.

Given a set of representative curve tokens in a curvature group, if the extracted curve terminates as an endpoint of a junction, we are done.

Suppose that the curve terminates with a smooth connection. Then the underlying smooth connection should be resulted by two connecting curves at *different* scales. Let  $\mathcal{S}_1$  and  $\mathcal{S}_2$  be the corresponding representative curve groups. In other words, smooth scale transition occurs between  $\mathcal{S}_1$  and  $\mathcal{S}_2$ .

Because we aim at preserving as much details as possible, to handle the scale transition, an *extrapolated* curve for  $\mathcal{S}_1$  is produced by tensor voting if the scale of  $\mathcal{S}_1$  is smaller than that of  $\mathcal{S}_2$ , after including and resampling the curve tokens in  $\mathcal{S}_2$  that are close to  $\mathcal{S}_1$ .

By using the stick voting field, if the extrapolated curve derived from  $\mathcal{S}_1$  can form a smooth connection with the curve derived from  $\mathcal{S}_2$ , then we are done (Fig. 5(a)). Else, the discontinuity (in the form of an endpoint or a junction) will be localized readily because the relevant curve features have been extracted or extrapolated (Fig. 5(b)).

### 2.3.2 Region extraction

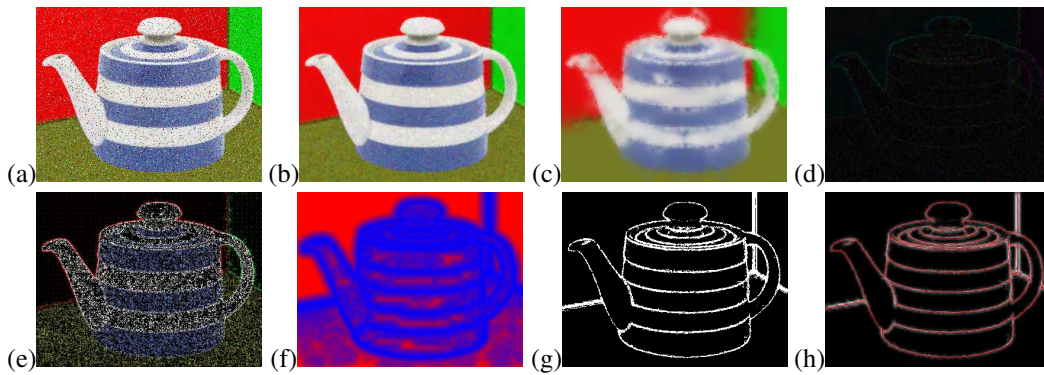
Given the representative density groups with their respective optimal scales (field sizes) obtained after extracting representative region tokens, region reconstruction proceeds from large-scale (low-density) region to small-scale region. For each density group, tensor voting at the optimal scale is performed to generate point tokens according to the local density at each representative token. The point generation is constrained by the region boundary curves extracted in the previous subsection.

Shown in Fig. 2(c) are the region tokens which are used to extract the region boundary curves. Depicted in Fig. 2(d) are the extracted curves where smooth scale transition occurs or junctions are localized. These curves and junctions are instrumental in constraining the following color generation step in order to preserve the image discontinuities.

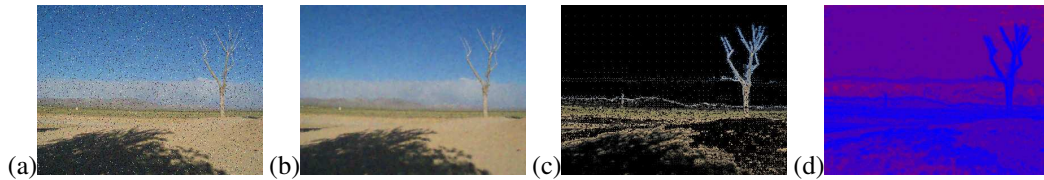
Finally, the color generation step for determining the colors of all pixels is proceeded as follows. For each pixel:

1. A representative token is included as the neighboring tokens to the pixel if a) the pixel is within the neighborhood of the token and b) the line joining the pixel and the token does not intersect with any extracted curves. The neighborhood of the token is defined by the optimal scale at each token. b) is used to prevent the mixture of the colors in two regions.
2. The color of the pixel is given by the weighted average of the colors of the neighboring tokens, where the weights of each token is given by its optimal scale.

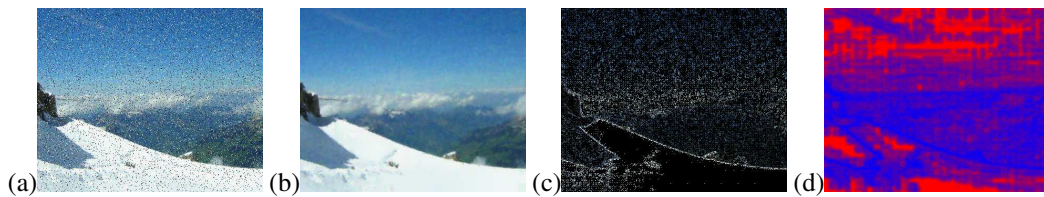
Fig. 2(a)–(b) shows the input image and the reconstructed image based on the minimal set of representative tokens shown in (e). Fig. 2(c) shows the result of color generation without curve constraints. The image difference between (b) and the noiseless input are shown in Fig. 2(d). Two additional reconstruction results based on the extracted representative tokens are shown in Fig. 3 and Fig. 4, where real images are tested. The size of the input image and the



**Figure 2.** (a) Noisy input image (size is 331KB). (b) Reconstructed image. (c) Reconstructed image without region boundary extraction. Colors in different regions are mixed. (d) Image difference between the noiseless image and (b). (e) Representative tokens (size is 88.9KB). (f) Optimal scales in color codes. (g) Region boundary tokens extracted after multiscale analysis. (h) Region boundary extracted. Precise curve position is detected by maximum vote saliency along curve normal.



**Figure 3.** (a) Noisy input image (size is 351KB). (b) Reconstructed image. (c) Minimal set of representative tokens (size 65.6KB). (d) Color-coded optimal scales. Note the tree trunk and desert are of different scales.



**Figure 4.** (a) Noisy input image (size is 441KB). (b) Reconstructed image. (c) Minimal set of representative tokens (size is 64.1KB). (d) Color-coded optimal scales.

representative tokens from which the faithful reconstruction or denoised image are compared.

### 3 Conclusion

In this paper, we propose to apply multiscale tensor voting to image denoising and compression in the context of perceptual grouping. While more experimental validation is needed, our results do show the promise of the multiscale reconstruction approach. The three major components of our multiscale scheme are: token grouping for rejecting noises and labeling the input tokens into the feature types detected by tensor voting; representative token extraction for deriving the minimal number of tokens for faithful reconstruction subject to the sampling theory and the smoothness of tensor voting fields; and multiscale feature extraction for extracting multiscale features integrated by smooth scale transitions or discontinuities. We have shown the efficacy of our method by performing multiscale reconstruction using real images, with application to image denoising and

compression.

### References

- [1] T. Lindeberg. Scale-space: A framework for handling image structures at multiple scales. In *CERN School of Computing*, pages 695–702, 1996.
- [2] T. Lindeberg. Principles for automatic scale selection. *Handbook on Computer Vision and Applications*, 2:239–274, 1999.
- [3] G. Medioni, M.S. Lee, and C.K. Tang. *A Computational Framework for Segmentation and Grouping*. Elsevier, 2000.
- [4] G. Papandreou and P. Maragos. A cross-validators statistical approach to scale selection for image denoising by nonlinear diffusion. In *CVPR05*, pages I: 625–630, 2005.
- [5] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE TPAMI*, 12(7):629–639, 1990.