

Coding Techniques for Oversampled Steerable Transforms

Baltasar Beferull-Lozano and Antonio Ortega
Department of Electrical Engineering-Systems,
University of Southern California,
Los Angeles, California 90089-2564.

Phone: (213)740-4679, Fax: (213) 740-4651

E-mail: beferull@sipi.usc.edu and ortega@sipi.usc.edu

Abstract

In this paper we study signal representation using oversampled steerable transforms. While in general it may not be efficient to use an oversampled representation for applications like compression, our work investigates efficient techniques for representing the oversampled data, given that after oversampling there exists substantial redundancy. We discuss different strategies which take advantage of this oversampling by establishing some consistency constraints on the representation that reduce uncertainty in the quantization. This results in a coding gain as we increase the oversampling in the steerable transform (number of orientations). Thus, while in general it will not be possible to achieve as good compression performance as with a critically sampled transform, having a compressed steerable representation will be useful for applications where feature is needed (many significant image features can be extracted from an orientation analysis), and where for performance reasons it is preferable not to have to decompress and analyze each image (as may be necessary if standard non-steerable transforms are used for compression.)

1. Introduction and Motivation

There are many different features that can be extracted from an orientation analysis in an image. This has been proved to be very useful in many different vision and image processing tasks, such as segmentation or texture analysis [3, 6]. The steerable transform allows us to analyze an image in any arbitrary orientation with a certain angular bandwidth resolution, but its main drawback is that its powerful multipurpose representation comes together with a high redundancy, which makes steerable transforms inferior to regular wavelets in terms of compression efficiency. The images in a data base are normally compressed using

either a DCT or wavelet based algorithm, because of their critical sampling as opposed to overcomplete decompositions. However, orthogonal and biorthogonal wavelet transforms, like all of the linear transforms commonly used in image compression, have significant drawbacks in the representations they produce, such as lack of shift and rotation invariance and no selectivity of orientation. This lack of selectivity in orientation in these critically sampled transforms, limits the characterizations of an image that can be obtained from their coefficients. Steerable transforms overcome all these disadvantages while keeping all the properties of traditional subband transforms, except critical sampling. Therefore, in the context of feature extraction over many different compressed images, we would have to first decompress each image and then apply a steerable transform to the decompressed image. Using directly a steerable transform to code the images, we can extract all these many different features directly in the transformed domain.

In this paper, we study angular oversampling in the context of steerable transforms (which were not considered by prior research in this area) and explore techniques to represent efficiently this oversampled data. The angular oversampling or oversteering is also motivated because it allows us to establish some “consistency” constraints [1, 4, 5] on the coefficients of a steerable representation with many orientations (oversteered representation), which reduces the amount of information lost in the quantization process and thus increases the accuracy and resolution of the corresponding coefficients.

We first give a brief overview of multiscale steerable transforms and then we analyze the relevant properties of steerable transforms that will be used to define formally the consistency constraints. Next, we explain different ways to represent the oversampled data placing an emphasis on the advantages of coding when we increase the number of orientations in the steerable transform (oversteering). We also explain how to represent efficiently the oversampled data

when we have oversteering and how the consistency constraints can be used to reduce uncertainty in the quantization process. Finally, we provide some preliminary experimental results that support our approach.

2. Overview of Steerable Transforms

A filter (or function) $h(x, y)$ is called steerable if transformed versions of this filter can be expressed using linear combinations of a fixed set of basis filters. If the transformation is a translation, then the filter is said to be shiftable or steerable in position; if the transformation is a rotation, then the filter is said to be steerable in orientation or commonly steerable and the basis filters are normally called steerable basis filters. Given a set of steerable basis filters, we can apply them to an image and since convolution is linear, we can interpolate exactly, from the basic responses of the basis filters, the output of a filter tuned to any orientation we desire, without explicitly designing and applying explicitly different filters for each one of the desired orientations. Following the analytical approach to steerability developed mainly by Freeman and Adelson [3] and later extended by Simoncelli [2], one can design filters, with certain restricted analytical forms, which are polar separable in the Fourier domain and where the steerable basis filters are steered versions of the steerable filter itself. Thus, if a filter $H(\omega_x, \omega_y)$ is polar separable in the Fourier domain, then $H(\omega_x, \omega_y) = B(r)G(\phi)$, and then steerability of $H(\omega_x, \omega_y)$ is equivalent to shiftable of $G(\phi)$. Given $G(\phi)$, one can find a set of N angles $\phi_0, \phi_1, \dots, \phi_{N-1}$ and a set of N interpolation functions $\{b_0(\phi), b_1(\phi), \dots, b_{N-1}(\phi)\}$ such that the following is satisfied:

$$G(\phi - \alpha) = \sum_{n=0}^{N-1} b_n(\alpha)G(\phi - \phi_n) \quad \forall \alpha \quad (1)$$

Also, we can impose a recursive pyramidal structure on the radial part, $B(r)$, so that we can construct a multi-scale, self-inverting and overcomplete pyramid decomposition (tight frame), where we have the possibility to interpolate exactly, at every scale independently, any subband steered to any desired orientation. The resulting transform is overcomplete by a factor of approximately $\frac{4N}{3}$, where N is the number of basis steerable filters.

3. Relevant Properties of Oversampled Steerable Transforms

In a steerable transform with N basic orientations, we have potentially an infinite dictionary at every scale or level, since given the N coefficients $\{c(\phi_0), c(\phi_1), \dots, c(\phi_{N-1})\}$, any coefficient (orientation)

(for the same spatial location) will be given as in (1), that is:

$$c(\phi) = \sum_{n=0}^{N-1} b_n(\phi)c(\phi_n) \quad \forall \phi \quad (2)$$

and we can determine the curve of the values of all the coefficients for the interval of angles $[0 \ \pi]$. This implies that if we are given 4 quantization intervals at the 4 basic orientations (angles), there is a certain deterministic region of uncertainty at any other angle. We will discuss later how to calculate it. Since only N basic orientations are enough to interpolate exactly any other orientation, if we have $M > N$ quantization intervals, the steerability constraint induces a ‘‘consistency’’ constraint on these intervals which will help to reduce uncertainty in some of the quantization intervals. In this paper, we use a steerable transform with $N = 4$ basic orientations, which corresponds to an angular Fourier response $G(\phi) = \cos^3(\phi)$ [2]. Given a chosen number N of basic orientations, depending on the basic orientations that are used, the interpolations functions are obtained to satisfy the steerability property [2]. As we will see later this has implications in the quantization process. Since we are interested in the oversteering case (i.e., having many orientations), we will generate, from the $N = 4$ basic orientations, a certain number $M > 4$ of subbands corresponding to M orientations.

4. Techniques to represent efficiently the oversampled data

All the discussion that follows concentrates on the problem of coding all the subbands (orientations) in only one level. This is the key problem in these types of oversampled data due to the large number of orientations that are present, and more specifically it focuses only on the problem of removing redundancies between different subbands. This particular coding problem can be solved efficiently by using consistency constraints and it makes a difference between a steerable transform and any other regular wavelet transform. On top of this, we could use any zero-tree based algorithm, to remove statistical dependencies across the scales, and also any context-based coding algorithm to remove redundancy inside each individual subband, all of them already developed in the context of wavelet transforms. It is clear that it is not useful at all to code each subband independently because of the large correlation among different subbands. There are basically two different ways to represent the oversampled data making explicit use of the correlation among different subbands. We explain these two approaches in the following subsections.

4.1. Selection of maximums

An important fact we have observed which is related to the steerable representation is that angular oversampling permits to localize most of the image energy in a few coefficients (energy compaction in angle). We have generated, in a 1 level pyramid, from the 4 basic orientations, 10 and 100 equally spaced (in angle) different orientations from 0 to π . Then, for each number of orientations, we have selected at every spatial location, the maximum coefficient in magnitude out of all the orientations and set to 0 the rest of coefficients (corresponding to the other orientations), and then, we have performed a simple thresholding in magnitude over the previously selected maximums. Finally, we reconstruct the original image (“Lena”) from the resulting coefficients. In the figure 1, we see that we localize energy more efficiently when we increase the number of orientations, so there is a substantial gain in energy compaction. This energy compaction in angle motivates the following

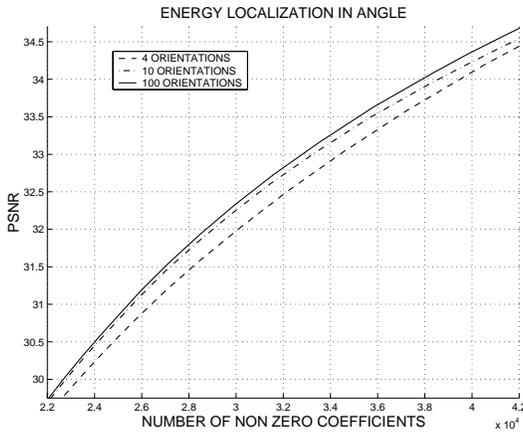


Figure 1. Energy compaction in angle

method to represent the data:

1. For every spatial location, depending on the coding accuracy we want to achieve, we select the T ($1 \leq T \leq 4$) coefficients (angles) with maximum magnitude and quantize their value according to some scalar quantizer whose stepsize will depend also on the coding resolution we want to have.
2. For every spatial location, using these T quantized maximums and other $4 - T$ angles which are assumed to have value 0, we can estimate the steerable curve and therefore we can predict the values of the coefficients that correspond to the other $M - 4$ angles or orientations.
3. Using the predicted values, we quantize the prediction error of the other $M - 4$ coefficients.

This method, however, has the disadvantage that the coding resolution is only based on the increase in the energy localization as we increase the number of orientations, but it is not taking full advantage of the properties of smoothness of the steerable curve.

4.2. Use of Consistency Constraints

In this case, we establish two constraints, one due to the smoothness (steerability) constraints on the steerable curve linking all the angular coefficients of all the subbands (for every spatial location) and another one due to the quantization itself. As we increase the oversampling, we will have a reduction in the reconstruction errors (improving coding accuracy) which will try to compensate the increase in the bit rate. We explain this technique, which is our fundamental approach, in the following section.

5. Consistency and Quantization

There are two ways to take advantage of the consistency constraints originated from the steerability property.

The first strategy is supported by POCS theory [7]. Since the steerable curve belongs to the space spanned by the interpolation functions $\{b_0(\phi), b_1(\phi), b_2(\phi), b_3(\phi)\}$, and we also know the quantization interval each coefficient belongs to, we can iterate projections on the following 2 sets: the set of all functions spanned by the interpolation functions, and the hypercube defined by the M quantization intervals. It is straightforward to see that both sets are convex and POCS theory tells us that each time we iterate we get a reconstruction value for each coefficient which is closer to the intersection of both sets (which will correspond to the original non quantized coefficient). After a sufficient number of iterations, we then perform differential entropy coding, that is, we do entropy coding on the quantized differences (indexes) between coefficients of adjoining angles. The total number of bits can be estimated by multiplying these entropies by the corresponding number of coefficients. This differential entropy coding is motivated from the fact that the steerable curve is always a smooth curve because of the smoothness of the interpolation functions. This approach, however, may not be practical because of its very high computational complexity, since the projection on the set of interpolation functions is carried out through a least squares fitting which is prohibitive in any coding algorithm.

In the second strategy, using 4 quantization intervals at 4 basic angles, we make use of steerability to constrain the region where all the $M - 4$ coefficients should fall in (region of uncertainty). We can use any 4 angles, and each group of 4 angles will give rise to a different region of uncertainty. The intersection of all these regions of uncertainty will determine another region of uncertainty that will tend to be

smaller due to the correlation among the different orientation subbands. We begin first quantizing the first 4 basic coefficients and then we can repeat this process each time we add a new orientation, by considering new different sets of 4 angles. Actually, this can be seen as being equivalent conceptually to the least squares fitting. The fitting with the interpolation functions tells us approximately where the original non quantized coefficients of the different angles should be. This same information can be obtained by performing the intersections of all the different regions of uncertainty that come from the different groups of 4 angles that we have as we keep adding more quantization intervals at more angles, up until we have M quantization intervals and we have performed all the possible intersections, which means that we cannot reduce more the uncertainty of the quantization intervals. Then, the reconstruction values will be taken as the middle points of the final quantization intervals. Notice that given a chosen number M of total orientations, we can precalculate easily the values of all the M sampled values of each of the interpolation functions that we need by using the linear equations described in [2] for general groups of 4 angles. Once we have these interpolation values, the complexity of calculating the intersections is very small as compared to a least squares fitting, because it only requires a few comparisons. In the following section, we solve the key problem of finding regions of uncertainty.

5.1. Calculating Regions of Uncertainty

We will call $R_U(\phi)$ and $R_L(\phi)$ the upper and lower limits of the region of uncertainty. These upper and lower limits are defined for every angle ϕ , so they will be curves themselves, which will bound a certain region of uncertainty that we will call $R(\phi)$. This problem can be stated as a linear programming problem for every angle ϕ , in the following way: given any 4 quantization intervals $\{I(\phi_0), I(\phi_1), I(\phi_2), I(\phi_3)\}$ we have to find the values $\{c(\phi_0), c(\phi_1), c(\phi_2), c(\phi_3)\}$ which are the solution of the following two following linear programming problems:

$$R_U(\phi) = \max \sum_{n=0}^{n=3} c(\phi_n) b_n(\phi)$$

$$R_L(\phi) = \min \sum_{n=0}^{n=3} c(\phi_n) b_n(\phi)$$

such that for $n = 1, 2, 3, 4$ we have that $l_n \leq c(\phi_n) \leq u_n$ and $\Delta_n = u_n - l_n$ are the stepsizes; the width of the region of uncertainty $W(\phi)$ is just $W(\phi) = R_U(\phi) - R_L(\phi)$. We notice that these linear programming problems have bounded solutions (constrained set is bounded) and that the solutions will be given for values of $\{c(\phi_0), c(\phi_1), c(\phi_2), c(\phi_3)\}$ which will correspond

to borders (upper or lower) of the quantization intervals $\{I(\phi_0), I(\phi_1), I(\phi_2), I(\phi_3)\}$. The solution of this problem is easily found to be the following:

$$R_U(\phi) = \sum_{n=0}^{n=3} U(\phi_n) b_n(\phi) \quad U(\phi_n) = \begin{cases} u_n & \text{if } b_n(\phi) > 0 \\ l_n & \text{if } b_n(\phi) < 0 \end{cases}$$

$$R_L(\phi) = \sum_{n=0}^{n=3} L(\phi_n) b_n(\phi) \quad L(\phi_n) = \begin{cases} l_n & \text{if } b_n(\phi) > 0 \\ u_n & \text{if } b_n(\phi) < 0 \end{cases}$$

The figures 2 and 3 show the region of uncertainty and the width of this region for a particular case where the basic quantization intervals have an stepsize of 5 and the basic angles are $0, \frac{\pi}{4}, \frac{\pi}{2}$ and $\frac{3\pi}{4}$. It can be seen in these figures that the width of uncertainty increases as we move away from the basic angles which is completely logical.

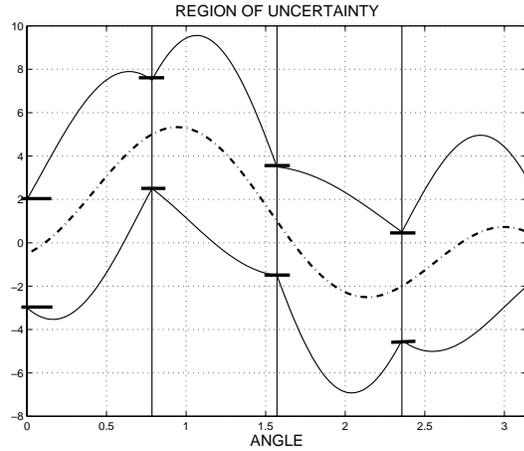


Figure 2. Example of a region of uncertainty

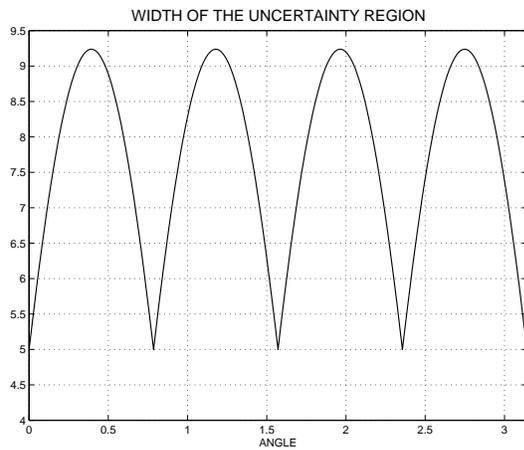


Figure 3. Example of the width of uncertainty

The solutions of these linear programming problems have the following interesting properties:

1. The width of the uncertainty region does not depend on the relative position of the quantization interval centers but it only depends on the 4 basic angles that are used and the widths of the quantization intervals.
2. The central curve of the resulting region of uncertainty satisfies steerability, that is, $\hat{c}(\phi) = \sum_{n=0}^3 \hat{c}(\phi_n) b_n(\phi)$ where $\hat{c}(\phi_n)$ is the middle point of the quantization interval $I(\phi_n)$.
3. These 2 properties imply immediately that the linear programming problem can be normalized to the one with all the intervals centered at 0, where the upper and lower bounds are symmetric.

This allows us to precompute, for different groups of 4 angles, the widths of the uncertainty regions as functions that depend only on the stepsizes and these functions will be valid independently of the quantization values.

5.2. Experimental Results

We have performed some experimental results with the “Lena” image and in particular we have studied the coding performance of our algorithms on the 3rd level of the steerable pyramid which is supposed to have important information because of its low frequency contents. As we mentioned before we have used a steerable transform with 4 basic orientations and we have oversteered the representation to 8, 16 and 32 orientations. We have compared the coding performance between the non oversteered case, that is, only 4 basic orientations (with direct quantization) chosen equally spaced as $0, \frac{\pi}{4}, \frac{\pi}{2}$ and $\frac{3\pi}{4}$, and the 3 other cases with oversteering. The comparison has been made in terms of the total number of bits which is measured as explained in section 5 and the MSE is averaged over all the orientations that we have in each case and thus, it is measured on the transformed domain. In order to get a range of values for the MSE, we have changed the value of the stepsize for the scalar quantizer which is applied initially at every angle. Figure 4 shows how the MSE is reduced as we increase the oversampling. From 4 orientations to 8 and 16 orientations, the improvements are 2.5 dB and 6 dB approximately, respectively. We also notice that for 32 orientations, the MSE is almost not reduced with respect to 16 orientations. The interpretation of this is that the smoothness of the steerable curve makes useless to use more constraints. In figure 5, we can see that there is a gain for low rates in the case of using 16 and 32 orientations with respect to using only the 4 basic orientations. This is because for these low rates, the reduction in MSE as we increase the number of orientations is faster than the corresponding increase in the number of bits, which results in a coding gain.

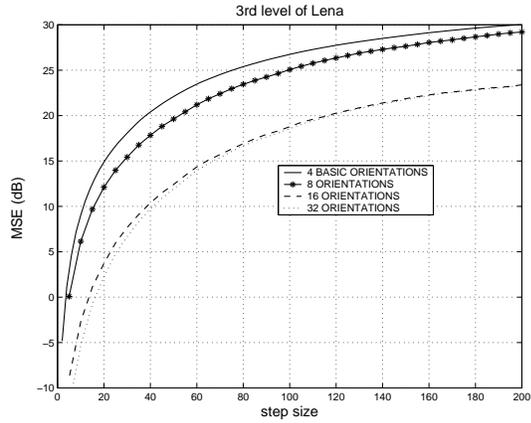


Figure 4. MSE vs stepsize

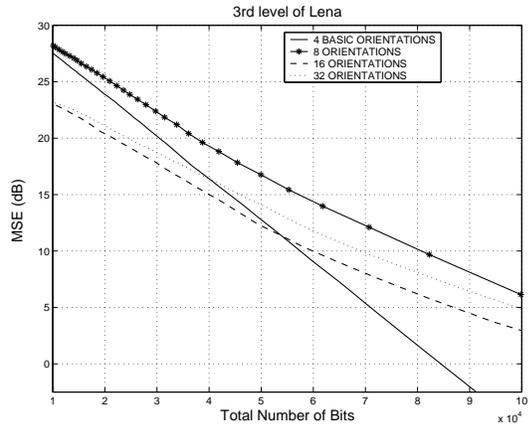


Figure 5. MSE vs the total number of bits

References

- [1] Z. Cvetkovic. Source coding with quantized redundant expansions: accuracy and reconstruction. *Data Comp. Conf.*, pag. 344–353, 1999.
- [2] E. Simoncelli, W. Freeman and D. Heeger. Shiftable multi-scale transforms. *IEEE T. on Inf. Th.*, 38(2):587–602, 1992.
- [3] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE T. on Pat. Anal. and Mach. Intell.*, 13(9):891–906, 1991.
- [4] V. Goyal. Quantized overcompleted expansions in R^n : Analysis, synthesis, and algorithms. *IEEE T. on Inf. Th.*, (1):16–31, 1998.
- [5] N.T. Thao and M. Vetterli. Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates. *IEEE T. on Sig. Proc.*, 42(3):519–531, 1994.
- [6] E. Simoncelli and J. Portilla. Texture characterization via joint statistics of wavelet coefficient magnitudes. *Int. Conf. on Image Proc.*, 1998.
- [7] D. C. Youla. *Mathematical theory of image restoration: the method of convex projections*. Image Recovery Theory and Application, 1987.