SYMMETRIC RESIDUE PYRAMIDS - AN EXTENSION OF BURT LAPLACIAN PYRAMIDS

Subrata Rakshit and Malay K. Nema

Centre for AI and Robotics (CAIR)
Raj Bhavan Circle, High Grounds
Bangalore 560001, India
subrata@cair.res.in, malay@cair.res.in

ABSTRACT

The Laplacian pyramid was proposed as a compact image code by Burt. However, it is shown that it does not achieve optimal signal compaction due to significant overlap between subbands in the frequency domain. In addition, it also leads to a slightly redundant signal representation. An extension of the Laplacian pyramid is proposed that exploits this redundancy to achieve better signal compaction. In contrast to the unidirectional fine to coarse approach used by all multiscale representations, the proposed signal representation adopts a bi-directional approach. Each subband finally depends on both its coarser and finer neighboring subbands. The representation is symmetric in the sense that each subband is a product of an approximation process (based on a finer approximation) as well as a prediction process (based on a coarser approximation). Applications are in the areas of image characterisation, zooming and compression.

1. INTRODUCTION

The Laplacian pyramid proposed was proposed by Burt in [1]. It was presented as a compact code, in that most of the originial signal energy would be encoded in a few of the transform coefficients. The Laplacian pyramid was also a multiresolution representation of the signal and one could construct an associated Gaussian pyramid which essentially consisted of a series of fine to coarse approximations of the input. Subsequently, the even more compact wavelet representation gained in popularity due to its mathematical elegance. A key difference between the two is that the Laplacian pyramids are an overcomplete (redundant) representation unlike the wavelets which are just complete. The redundancy in Laplacian pyramids ($< 33\%$ for 2D) gives a lot of flexibility in filter design. The Laplacian pyramid gives exact reconstruction for arbitrary filters, though the subbands and multiscale approximations are useful only for reasonably designed lowpass filters. The essential similarity between Laplacian pyramids and wavelets is the idea of forming successive approximations of an input and forming subbands based on the differences in these approximations. The process is entirely bottom-up with each subband depending solely on the approximation at its own level.

The subbands as produced above are not in fact independent. This arises in the case of wavelets because high frequencies are often caused by sharp edges. As a result, these frequencies are predictably coherent and knowledge of some of these components allows one to guess the others. The predictability is even greater in Laplacian pyramids. The overcompleteness of the representation forces certain relationships among coefficients across subbands. Also, the subbands have a lot of overlap in the frequency domain and hence are correlated. In this paper we present a way to exploit the redundancy to remove this correlation between Laplacian subbands. The redefined subbands are now the product of an approximation process as well as a prediction process. This symmetric residue (SR) pyramid can be used to perfectly construct the Laplacian pyramid and therefore the original signal. This results in a signal representation that achieves better energy compaction with more independent subbands. As a side benefit, it gives a novel way to double image sizes by generating coherent high frequencies.

Research supported entirely by CAIR(DRDO).

2. INTERBAND PREDICTION IN PYRAMIDS

The Laplacian pyramid defined by Burt [1] is constructed as follows. Given an approximation at level i as G_i, the next approximation is $G_{i+1} = ss(lpf(G_i))$ where lpf is a low pass filtering operation and ss is a subsampling operation. The bandpass Laplacian subbands are generated as $L_i = G_i - exp(G_{i+1})$ where exp is an up-sampling operation followed by smoothing that exapnds the image. As the bandpass Laplacians, L_i, are not subsampled the representation is overcomplete. G_i can be exactly reconstructed from L_i and G_{i+1} as $G_i = L_i + exp(G_{i+1})$ by construction.

Corresponding to an input image with N^2 pixels, the Laplacian pyramid representation has $4 * N^2/3$ pixels. Clearly, this implies there are many more possible pyramids corresponding to N^2 pixel images than there are N^2 pixel images. It has been shown in [2] that this fact can be exploited to use the Laplacian pyramid as an error correcting code for images. Let $P = \mathcal{P}(I)$ denote the process of constructing a pyramid P from an image I and let $I = \mathcal{I}(P)$ denote the process of reconstructing an image I from a pyramid P. Error correction is based on the fact that $P = \mathcal{P}(\mathcal{I}(P))$ if and only if $P = \mathcal{P}(I)$ for some image I [2]. A variant of this idea can be used for interband prediction as well.

Given a G_{i+1}, not all possible L_i are candidate Laplacians at level i. The true L_i will satisfy the consistency condition

$$G_{i+1} \quad = \quad ss(lpf(L_i + exp(G_{i+1}))). \qquad (1)$$

This relationship does not specify a unique L_i for every G_{i+1}. In order to see this, consider the mapping from G_i to G_{i+1}. Due to the subsampling operation, G_{i+1} has only 1/4 as many pixels as G_i. Thus multiple G_i map on to the same G_{i+1}. More precisely, let there be K different G_i and P different G_{i+1}. Since L_i have as many pixels as G_i, there will be K different L_i as well. Let Q be the number of G_i such that their $G_{i+1} = \mathbf{0}$. Due to linearity, these G_i can be added to any other G_i without changing the corresponding G_{i+1}. (These G_i contain all possible high frequency patterns that get completely blocked by the lpf operation.) Conversely, if two different G_i map to the same G_{i+1}, then the G_{i+1} corresponding to the difference of these G_i must be **0**. Thus each G_{i+1} has associated with it exactly Q different G_i which implies that $Q.P = K$. Correspondingly, there will be Q different L_i that satisfy Eqn 1 for each G_{i+1}.

Let us analyze the above in information theoretic terms. If G_i and L_i have n^2 pixels, then G_{i+1} will have $n^2/4$ pixels. Since G_i completely specifies L_i and G_{i+1}, the $5n^2/4$ pixels of G_{i+1} and L_i must be coding only as much information as the n^2 pixels of G_i. Since G_{i+1} is unconstrained, the n^2 pixels of L_i must be having only $3n^2/4$ effectively independent pixels once the condition of consistency is imposed. The information that can be encoded in images is proportional to the number of pixels and hence varies as the logarithm of the number of possible distinct images. By the notation used above, we then expect $\log(Q) = \log(K) - \log(P)$ which is identical to the relation $Q.P = K$ derived above. Hence the P different G_{i+1} partition the K different L_i into K/P cosets of size Q each. The exact values of K, P and Q depend on number of pixels, bits per pixel and the low pass filter, lpf. By the Slepian-Wolf theorem [3], the knowledge of G_{i+1} should lead to an efficient coding of L_i proportional to $\log(Q)$ rather than $\log(K)$.

In terms of a frequency domain analysis, the Laplacian subbands multiply encode certain parts of the input spectrum. The consistency requirement of Eqn 1 implies that the lower frequency components of L_i need to be consistent with similar components in G_{i+1}. As can be seen in Fig. 1, the frequency responses of the Laplacian pyramid subbands show significant overlap. In particular, each Laplacian subband has a significant 'tail' that picks up low frequency information from the next subband. This introduces dependencies in addition to that caused by the fact that sharp edges produce coherent harmonics at various resolutions. The information in each subband can be decomposed into three components. (i) The low frequency component that is present in higher subbands. (ii) High frequency components that are coherent with harmonics present in higher subbands. (iii) High frequencies that are independent of harmonics present in other subbands. Clearly, the first can be rigourously computed. Given a prior model of edges, the second may be predicted as per the model. The third component is the independent information of the subband and cannot be predicted.

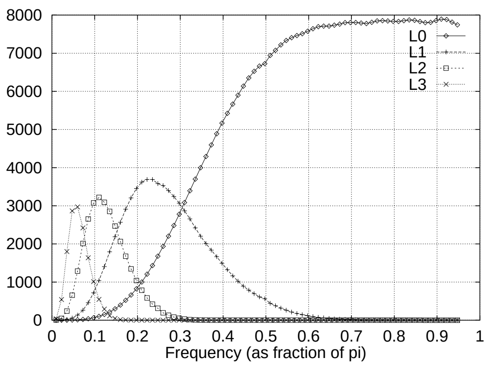

Fig. 1. Frequency responses of Burt Laplacian pyramid subbands

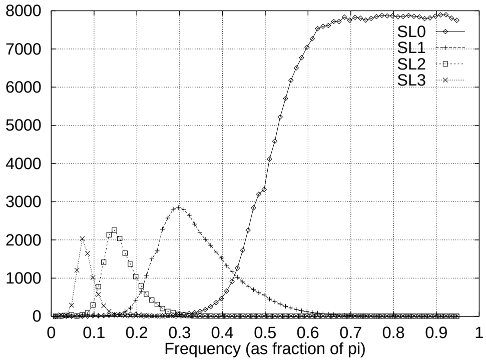

Fig. 2. Frequency responses of SR pyramid subbands

3. SYMMETRIC RESIDUE PYRAMIDS

An extension of the original Burt Laplacian pyramid is proposed that takes into account the predictability of the L_i subband based on G_{i+1}. The motivation for this extension is to achieve better signal compaction. In addition, we would like to emphasize that this compaction is being achieved by altering the way multiresolution representations are computed and not by designing a better filter or transform. Traditionally, the aim has always been to compute successive approximations and subband signal decorrelation was based on the orthogonality of the underlying basis functions (exactly in the case of wavelets, only approximately for pyramids). No attempt was made to formulate representations where subbands were mutually independent in an information theoretic sense. As a result, a cursory glance at a pyramid or wavelet representation shows a striking similarity across subbands and one can always tell subbands of one image from that of others. True signal independence is clearly a far cry. The presence of dependent low frequencies makes the Laplacian pyramid an obvious choice for this type of redundancy removal. However, model based prediction for coherent harmonics can be applied to wavelets as well. Models for such prediction and algorithms for image zooming based on this can be found in [4, 5].

Eqn 1 can be rewritten in terms of a consistency condition on the L_i itself. The L_i, when used to recreate a G_i from a G_{i+1}, must create a G_i that will again lead to the creation of the same L_i during pyramid creation process.

$$G_i = L_i + exp(G_{i+1}) \quad (2)$$
$$L_i = G_i - exp(ss(lpf(G_i))) \quad (3)$$

In practise, one does not know L_i and must start with a guess. An iterative process is used to converge to one of the many admissible L_i.

$$\hat{G}_i = L_i^j + exp(G_{i+1}) \quad (4)$$
$$\hat{L}_i = \hat{G}_i - exp(ss(lpf(\hat{G}_i))) \quad (5)$$
$$L_i^{j+1} = (L_i^j + \hat{L}_i)/2 \quad (6)$$

The process is stopped as soon as L_i^j has converged to one of the many solutions consistent with a given G_{i+1}, say L_i^*. This is determined by checking if $ss(lpf(\hat{G}_i)) \rightarrow G_{i+1}$. Beyond that the system may oscillate between various admissible L_i, picking up high frequency components. The estimated L_i^* can differ from the true, desired L_i only in terms of such high frequencies as would be totally blocked by the lpf operation. It is this difference pattern that constitutes the independent information that needs to be encoded by the subband at level i.

If one wishes to invoke a prior for L_0, then that can be done while choosing L_i^0 to start the iteration. In the absence of a prior, L_i^0 may be set to **0** or $exp(L_{i+1})$. Either ways, it starts without high frequencies and stays that way. Alternately, one could introduce selected high frequencies into L_i^0 to model (expected) sharp edges based on an analysis of G_{i+1}. The estimation of various types of priors is an ongoing activity and beyond the scope of this paper. The results reported here are for $L_i^0 = exp(L_{i+1})$ (no prior).

The Symmetric Residue (SR) pyramid may now be formulated as follows

1. Form the Burt Laplacian pyramid: $\{L_i, i = 0...n\}$, G_{n+1}
2. Recursively, compute the G_{i+1} and best predicted Laplacians L_i^*, $i = n...0$.
3. Define the SR pyramid subbands as $SL_i = L_i - L_i^*$
4. The SR pyramid is defined by $\{SL_i, i = 0..n\}$, G_{n+1}

The SR pyramid construction is a two pass algorithm. The first pass is a bottom up process leading to successive approximations. The second pass is a top down process based on inter-band prediction leading to subband independence. The creation of each subband now depends on the previous as well as next subband leading to a symmetric process. The information it encodes is the residual which is not encoded at the next subband or the previous one. The process is perfectly reversible as first the L_i and then the G_i can be recursively computed based on the SL_i and G_{n+1}.

Fig. 3. Two levels of Burt Laplacian pyramid for LENA. Left: L_0 subband, top right: L_1 subband and bottom right: G_2 subband

The frequency responses of the SR pyramid are shown in Fig. 2. As can be seen, in comparison with the responses for the Laplacian pyramid the SR pyramid subbands have much less overlap. In particular, the subbands do not have significant low frequency responses. The Laplacian pyramid and an SR pyramid for the LENA image is shown in Fig. 3 and Fig. 4 respectively. In the case of the SR pyramid, the subbands consist solely of ripples around sharp edges of the image. These are the residuals (corrections) that need to be applied to the predicted subbands. This is shown in detail in Fig. 5. The G_1 image for LENA is shown along with the L_0 subband that it predicts. This is quite close to the actual L_0 subband for LENA as shown in Fig 3. This leads to the expectation that compact encoding of L_i given L_i^* should be possible based on the Slepian-Wolf theorem [3].

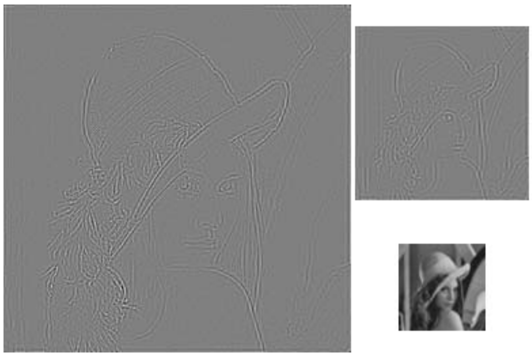

Fig. 4. Two levels of Symmetric Residue pyramid for LENA. Left: SL_0 subband, top right: SL_1 subband and bottom right: G_2 subband

Fig. 5. Based on the G_1 subband, the L_0 subband can be estimated. Compare the L_0 estimation in this figure to the true L_0 shown in Fig. 3. The difference is the SL_0 subband in Fig. 4

4. CHARACTERISATION OF SR PYRAMIDS

The SR pyramids are characterised by comparing some of their properties with those of the original Burt Laplacian pyramids. In order to carry out this exercise, the following set of images were used: LENA, MANDRIL, BARBARA, CLOWN and GIRL.

4.1. Energy Compaction

The ability of Laplacian pyramids and SR pyramids to encode images with most of the energy concentrated in the higher subbands is tested. The energy in each subband is measured as sum squared of all pixel values. The results are shown in Table 1 for 4 level pyramid representations. G_4, common to both, is not shown. The low values arise as pyramids are *not* snug (norm preserving) frames [2].

Band	Laplacian, L_i	SR, SL_i
Lev 0	1.426 %	0.756 %
Lev 1	0.298 %	0.093 %
Lev 2	0.109 %	0.031 %
Lev 3	0.042 %	0.014 %

Table 1. The energy in the subbands of Laplacian and SR pyramids as percentages of the input (G_0) energy. The results are averaged over five images mentioned in the text.

4.2. Independence of Subbands

The independence between subbands is measured in two ways. The first is based on correlation, wherein the smaller subbands are expanded to a common size and the cross correlation between these images is estimated. This is effective in determining if there is any signal component common between the two subbands. In addition, the mutual information between the two subbands is also estimated. This measure is based on the individual entropies and the cross entropy. (All entropies are calculated based on pixel intensity distributions only.) The mutual information tests for predicatbility between two signals.

Table 2 shows the extent to which subbands are correlated in Laplacian and SR pyramids. As can be seen, SR pyramid subbands are significantly de-correlated as compared to the Laplacian subbands. A similar trend is seen in Tab. 3 for the mutual information between subbands. As noted before, the SR pyramid results shown here are without the use of any priors. The use of priors is expected to improve these results.

	L_0	L_1	L_2	L_3
L_0	1.000	0.487	0.186	0.066
L_1	0.487	1.000	0.614	0.248
L_2	0.186	0.614	1.000	0.621
L_3	0.066	0.248	0.621	1.000

	SL_0	SL_1	SL_2	SL_3
SL_0	1.000	0.037	0.010	0.006
SL_1	0.037	1.000	0.052	0.019
SL_2	0.010	0.052	1.000	0.059
SL_3	0.006	0.019	0.059	1.000

Table 2. The top half shows the correlation between subbands in Laplacian pyramid. The bottom half shows correlation between subbands in SR pyramids. The SR pyramids subbands are more decorrelated. Results have been averaged over five images.

	L_0	L_1	L_2	L_3
L_0	5.502	0.371	0.183	0.173
L_1	0.371	5.401	0.940	0.427
L_2	0.183	0.940	5.758	1.175
L_3	0.173	0.427	1.175	6.025

	SL_0	SL_1	SL_2	SL_3
SL_0	5.097	0.103	0.064	0.067
SL_1	0.103	4.654	0.841	0.302
SL_2	0.064	0.841	4.913	0.711
SL_3	0.067	0.302	0.711	5.241

Table 3. The top half shows the mutual information (in bits) between subbands in Laplacian pyramid. The bottom half shows mutual information between subbands in SR pyramids. The SR pyramids subbands are more independent. The diagonal entries indicate the entropies of the histograms of the subbands. Results have been averaged over five images.

4.3. Image Interpolation

The SR pyramid formulation allows good interpolation and magnification of images. If a given image is considered as the G_0 approximation, then the next higher resolution would be the G_{-1} approximation. It may be computed as $G_{-1} = exp(G_0) + L_{-1}$ with $L_{-1} = \mathbf{0}$ for the Burt Laplacian pyramid. Similarly, one may use SR pyramids with $SL_{-1} = \mathbf{0}$ and wavelets with $H_{-1} = V_{-1} = D_{-1} = \mathbf{0}$. A comparison was performed based on the task of estimating the G_0 image from the G_1 image as this provided ground truth for SNR estimation. The Burt, SR and wavelet methods all used 5-tap filters (Daub $N = 2$ for wavelets). Results for interpolation using a windowed 21-tap sinc filter and nonlinear frequency domain interpolation (NFDI)[5] are also reported. The results shown below are an average over the five images mentioned earlier.

	Burt	Sinc21	NFDI	Wlet	SR pyr
SNR (dB)	18.51	18.11	18.93	17.45	21.21

5. CONCLUSION AND FUTURE WORK

The Burt Laplacian pyramid algorithm has been extended to include a top-down prediction process. The new representation, the Symmetric Residue pyramid, is also a complete representation capable of exactly reconstructing the original signal. The prediction process is shown to remove the redundancy in information encoding that is present in the original Laplacian pyramid. This is shown to lead to a representation that has better subband frequency characteristics, energy compaction properties and subband independence. The scheme allows for the use of priors to further reduce energy in subbands. The formulation and analysis of such priors is currently being carried out. One of the benefits of the prediction process is the ability to create higher resolution approximations that are consistent with a given image. Applications for compression and texture characterization will be pursued in future.

6. ACKNOWLEDGEMENTS

The authors wish to thank Director, CAIR, for his support and encouragement and Computer Vision group colleagues for their help.

7. REFERENCES

[1] Peter J. Burt, Edward H. Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans on Communications, vol 31, no 4, pp532-540. 1983.

[2] Subrata Rakshit, Charles H. Anderson, "Error correction with frames: the Burt Laplacian pyramid," IEEE Trans on Information Theory, vol 41, no 6, pp 2091-2093. 1995.

[3] D. Slepian, J.K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans on Information Theory, vol IT-19, pp471-480. 1973.

[4] Stephane Mallat, Sifen Zhong, "Characterization of Signals from Multiscale Edges," IEEE Trans on Pattern Analysis and Machine Learning, vol 14, no 7, pp 710-732. 1992.

[5] Hayit Greenspan, Charles H. Anderson, Sofia Akber, "Image Enhancement by Nonlinear Extrapolation in Frequency Space," IEEE Trans on Image Processing, vol 9, no 6, pp 1035-1048. 2000.