Computer Vision Winter Workshop 2010, Libor Špaček, Vojtěch Franc (eds.)
Nové Hrady, Czech Republic, February 3–5
Czech Pattern Recognition Society

Irregular Laplacian Graph Pyramid

Dan Shao and Walter G. Kropatsch

Vienna University of Technology,
Pattern Recognition and Image Processing Group,
Favoritenstr. 9/1832, A-1040 Vienna, AUSTRIA
shao@prip.tuwien.ac.at

Abstract *This paper presents a novel image representation, which incorporates the principles of Laplacian Pyramid into the irregular graph pyramid. The drawback of the regular Laplacian Pyramid is their lack to keep the topological structure of the image, due to the contraction process in building the Gaussian Pyramid. Irregular graph pyramid is able to hierarchically represent the topological structure of an image with multiresolution, where each level is a graph describing the image with various resolutions by contracting the graph from the level below. We build irregular Laplacian graph pyramid by storing the difference of levels in irregular graph pyramid. Experiments and results are presented in the paper to show the characteristic of the irregular Laplacian graph pyramid and some immediate advantages in computer vision applications.*

1 Introduction

Image pixels are in generally highly correlated, it is common to have several areas of an image sharing the same or similar pixel values. Therefore, it is redundant to encode the image information by each of its pixel values.

In order to design an efficient compression code, it is necessary to find a representation that decorrelates the image pixels.

Laplacian Pyramid is a versatile data structure with many attractive features for image preprocessing. It represents an image as a series of quasi-bandpassed images, each sampled at successively sparser densities [1].

The bases of the Laplacian Pyramid depend on the Gaussian Pyramid, in section 2 we will describe this process. Gaussian Pyramid's biggest drawback is their lack to keep the topological structure of the image, leading to inaccurate image structure from the original image over the contraction process. And regular Gaussian Pyramid is not shift invariant. Irregular graph pyramid overcomes this Gaussian Pyramid's drawback by keeping the topological structure of the image within the contraction process, and hierarchically represents the image with a parent-child relationship over all the pyramid levels.

The properties of the irregular graph pyramid and its applications in image processing motivated us to use the advantages of its topological/structural features in image encoding.

The main aim of this paper is to incorporate the regular Laplacian Pyramid concept into the irregular graph pyramid.

A highly relevant property of the Laplacian Pyramid, is to be able to have a progressive image transmission. In this type of progressive image transmission a coarse rendition of the image is sent first to give the receiver an early impression of image content, then subsequent transmission provides image detail of progressively finer resolution [2]. Due to the drawback of regular pyramid, the early impression of image content may show a different topology structure than the original image, for example, one connected component could be modified to appear as two connected components in coarse rendition, see Fig 1. With the irregular Laplacian graph pyramid, it is possible to solve this problem.

Major contributions of this paper are to incorporate the Laplacian Pyramid principles into the irregular graph pyramid, to overcome the major drawbacks of regular Laplacian Pyramid, the drawback of structure inconsistency.

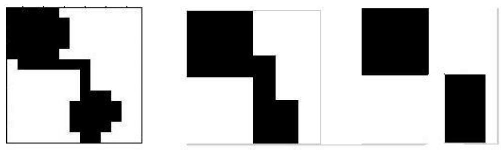

Figure 1: Fine to Coarse Images by Gaussian Pyramid

1.1 Organization of paper

In Section 2 we recall the regular Laplacian Pyramid. In Section 3 we explain the concept of combinatorial pyramid, which is pre-step of building irregular Laplacian Pyramid. In Section 4 we describe the method of building irregular Laplacian Pyramid, following by experiment and result discussion in section 5. In Section 6 we give the conclusion and open questions.

2 Recall of Laplacian Pyramid

Predictive coding is the base of Laplacian Pyramid, pixels are encoded sequentially in a raster format. However, prior to encoding each pixel, its value is predicted from previously coded pixels in the same way and proceding raster lines [1].

Laplacian Pyramid has been always related to Gaussian Pyramid, which is a low-pass filtered image sequence which is then subtracted from the original.

2.1 Gaussian Pyramid

The first step to build a Laplacian Pyramid is to low-pass filter the original image g_0 to obtain the first image level g_1. This g_1 image is decreased in resolution and sample density. In the same way g_2 is a reduced version of g_1. To build a Gaussian Pyramid this process continue from the reduced image, and so on.

Suppose the original image, which is the base of the pyramid, is represented as an array g_0 containing C columns and R rows of pixels. Each pixel represents the light intensity at the corresponding image coordinate by an integer I with values between 0 and K-1 [1].

g_1 from pyramid level 1 consists of a low-pass filtered version of the previous level, g_0 from level 0 . Where each value of the level 1 is computed as a weighted average of values from level 0 within a 5-by-5 window. The size of the weighting function is not critical [3]. In Fig. 2, is shown in 1D the reduction process, in 1D the density of the nodes are reduced by half while in 2D by fourth from level to level.

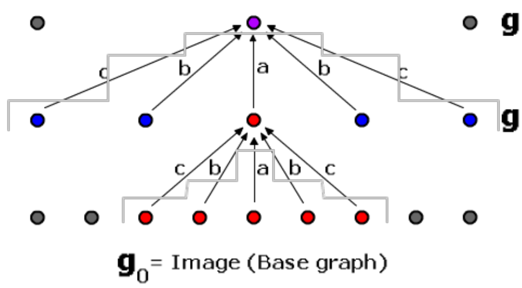

Figure 2: Gaussian Pyramid.

The reduction function from level to level is the averaging process such that, for levels $0 < l < N$ and nodes i, j, $0 \leq i < C_l$, $0 \leq j < R_l$.

Where N refers to the number of levels in the pyramid, C_l and R_l are the dimensions of the lth level.

$$g_l(i,j) = \sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n) g_{l-1}(2i+m, 2j+n) \quad (1)$$

The whole pyramid is only 4/3 the size of the original image. Each higher level of the pyramid is about half as large for each dimension as its previous level, as shown in Fig. 3.

The pyramid building process is equivalent to convolve the base image g_0 with a set of equivalent weighting functions h_l [1].

$g_1 = h_l \star g_0$

The effect of convolving an image with one of the equivalent weighting functions h_l is to blur, or low-pass filter, the image. The pyramid algorithm reduces the filter band limit by an octave from level to level, and reduces the sample interval by the same factor [1]. This is a very fast algorithm, requiring fewer computational steps to compute a set of filtered images than are required by the fast Fourier transform to compute a single filtered image [3].

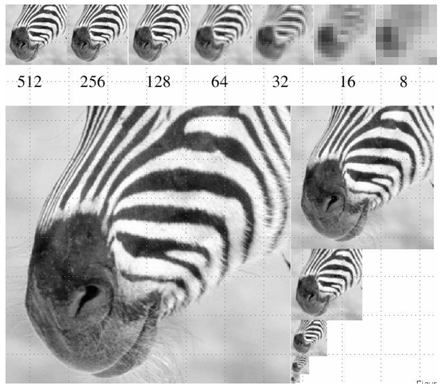

Figure 3: The levels of Gaussian Pyramid of dragon head (512x512), decreased half resolution by each level: 512x512, 256x256, 128x128, 64x64 and so on.

2.2 Laplacian Pyramid

The expansion function is the inverse of the reduction function.

It expands an $(M + 1)$ by $(N + 1)$ array into a $(2M + 1)$ by $(2N + 1)$ array by interpolating new node values between the given values of pixels.

The expansion applied to the array g_1 of the Gaussian Pyramid would result into an array $g_{l,1}$ which is the same size as g_{l-1}.

For levels $0 < l \leq N$ and $0 < n$ nodes $i, j, 0 \leq i < C_{l-n}$, $0 \leq j < R_{l-n}$

$$g_l, n(ij) = 4 \sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n) \bullet g_{l-1}\left(\frac{i-m}{2}, \frac{j-n}{2}\right) \quad (2)$$

If we apply expansion function l times to the image g_l, the result is $g_{l,l}$ which is the same size as the original image g_0.

The purpose of constructing the reduced image g_1 is that it will be used as a prediction for pixel values in the original image g_0.

To obtain a compressed representation, we encode the error image which remains when an expanded g_1 is subtracted from g_0. This image becomes the bottom level of the Laplacian Pyramid. The next level is generated by encoding g_1 in the same way.

The Laplacian Pyramid is a sequence of error images L_0, L_1, ..., L_{N-1}. Each is the difference between two levels of the Gaussian Pyramid. Thus, for $0 \leq 1 < N$, $L_l = g_l$ - EXPAND(g_{l+1}) [1].

In Laplacian Pyramid, the value at each node is the difference between the convolutions of two equivalent weighting functions h_l, h_{l+1} with the original image.

Gaussian Pyramid can be seen as a set of filtered copies of the original image, while the Laplacian Pyramid can be seen as a set of bandpass filtered copies of the image. Image features such as edges appear enhanced in the Laplacian Pyramid as show Fig. 9. These enhanced features depend on the size, where fine details are notable in $L_{0,0}$ and progressively vanish in the higher levels.

The original image can be recovered completely by expanding. First, expand L_N once and add it to L_{N-1}, then expand this image once and add it to L_{N-2}, and so on until level 0 is reached and g_0 is recovered [1].

3 Combinatorial Pyramid

An irregular graph pyramid combines graph structures with hierarchies. Each level is a graph representation describing the image with various resolutions by contracting the graph from the level below. Graph representations have been investigated widely during last decades for representing structural information in various domains in computer vision such as image segmentation and object recognition.

The graph structure is defined as follows: graph $G(V, E)$ consists of vertices $v \in V$and edges $e \in E$. An edge e connects two vertices,v, w, $e_i = (v, w)$. The vertices V and edges E of the image graphs carry not only the structural information, but also the additional information.

For this paper we consider using combinatorial maps to present the graph in each level of the irregular pyramid. A combinatorial map is a topological model which allows to represent subdivided objects as planar graphs. A 2D combinatorial map is defined by a triplet $M = (D, \sigma, \alpha)$ where D is a finite set of darts, σ is a permutation on D and α is an involution on D without fixed point [7]. For each dart, σ gives the next dart by turning around the vertex v in the positive orientation (clockwise); For each dart, α gives the other dart of the same edge e. There are always two darts corresponding to a same edge, α allows to retrieve edge e, and σ allows to retrieve the vertex v.

We can see in Fig. 4, the graph is an example of 2D combinatorial map.

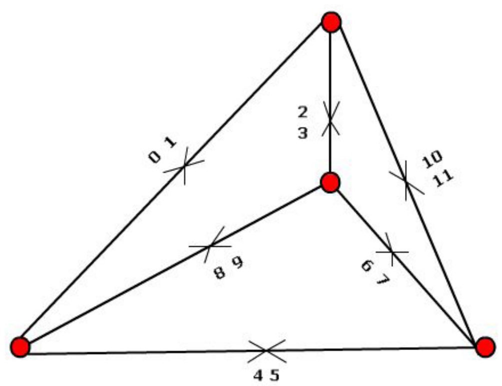

Figure 4: Combinatorial map, contraction/removal operations.

This map can be explicitly defined by giving the set of darts, permutations σ and involutions α.

D	0	1	2	3	4	5	6	7	8	9	10	11
σ	8	10	1	6	0	7	9	11	4	3	2	5
α	1	0	3	2	5	4	7	6	9	8	11	10

During the process of building up the irregular graph pyramid, contraction process removes the edges from the input graph while simultaneously merging together the vertices it used to connect [7]. And reduction process is to take the attributes of all children as input and then compute the parent's attribute as output.

Taking a simplified image of a cup as example (Fig 5), we build the base graph as the input image, where each vertex represent a pixel in the input image. Then use the contraction methods to build the irregular pyramid. Such approach would lead to a pyramidal structures like Fig 5:

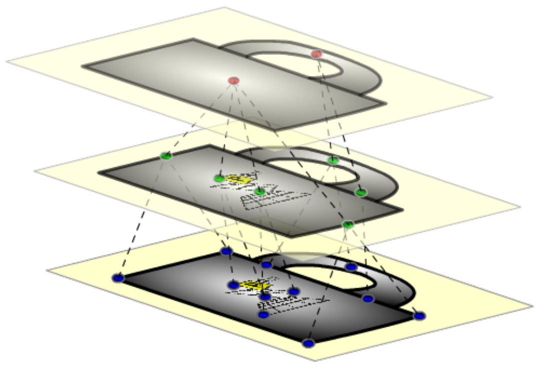

Figure 5: Irregular Graph Pyramid.

Level 0: The base level of the pyramid consists in a geometric description of the underlying image (here a simplified image of a cup).

Level 1: The second level of the pyramid, simpler boundaries are abstracted from base level (like the handle and the logo of the cup).

Level 2: Adjacent parts of the cup are grouped in order to represent compound abstract objects.

4 Irregular Laplacian Graph Pyramid

Similar as regular Laplacian image pyramid, the irregular Laplacian Pyramid stores the difference of the child's content with expanded content, but from irregular graph pyramid instead of Gaussian Pyramid. However, the expansion function is slightly different from regular Laplacian Pyramid. In this section, we will formalize the process of building the irregular Laplacian Pyramid.

4.1 Correction Vector

In the process of building up irregular graph pyramid, each level graph is obtained by contracting the graph from the level below. Parent inherits the position of the surviving child. The property of the parent is computed from the properties of all its children by a certain function, such as weighted average. Depending on the concrete task, this function varies. Let V_p refers the parent vertex, and V_c refers the child vertex:

$$pos(V_p) = pos(V_c) \quad \text{if } V_c \text{ survives}$$
$$d(V_c) = pos(V_c) - pos(V_p) \quad \text{if } V_c \text{ does not survive}$$

Each non-surviving vertex stores the difference between its coordinate and the its parent's coordinates. We call this difference as correction vector. Because the parent's position is equal to the surviving child's position, those surviving vertices have correction vector of value 0. As show in Fig 6,vertex V_{c_1} is the surviving vertex. So its parent V_p has the same

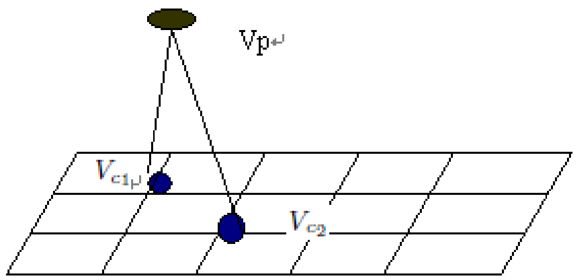

Figure 6: correction vector: Vertex V_{c_1} is the surviving vertex and vertex V_{c_2} is contracted

coordinate as V_{c_1}. The correction vector of surviving vertex V_{c_1} is 0. The correction vector of non-surviving vertex V_{c_2} is $\sqrt{2}$, which is the geometric distance of V_{c_1} and V_{c_2}.

4.2 Expansion Kernel

Expansion kernel expands (interpolates) the properties of the parents cells into the children's content at the higher resolution level. With correction vector, the position of the child can be obtained by adding its correction vector to its parent's position. G_k refers the graph in level k, and G_{k-1} refers to the graph in one level below.

$pos(V_c) = pos(V_p)$
where $V_p \in G_k, V_c \in G_{k-1} \cap G_k$
$pos(V_c) = pos(V_p) + L(V_c)$
where $V_p \in G_k, V_c \notin\in G_{k-1} \cap G_k$

Let's call the irregular graph pyramid as G, and the irregular Laplacian Pyramid as L. For the properties of vertices, each level of the Laplacian Pyramid can be obtain by taking the difference of the adjacent levels, same as in regular Laplacian Pyramid.

$$L_k = G_k - Expand(G_{k+1})$$

The reconstruction process is same as in the regular case. The original image can be reconstructed by expanding. First, expand G_{k+1} once and add it to L_k, then expand this image once and add it to L_{k-1}, and so on until level 0 is reached and G_0 is recovered.

$$G_k = Expand(G_{k+1}) + L_k$$

The main process of building irregular Laplacian Pyramid is divided by three steps: 1). Build Irregular Graph Pyramid G on the target image. 2). For each level k in the G, We expand level G_{k+1} using the correction vectors to have the same size of content as the level k in G_k. 3). Take the difference of the expanded level G_{k+1} with the level G_k.

5 Experiment

We take a gray image as example, to demonstrate the experiment of building the irregular Laplacian Pyramid on it. For initialization, we convert this gray image into a graph which

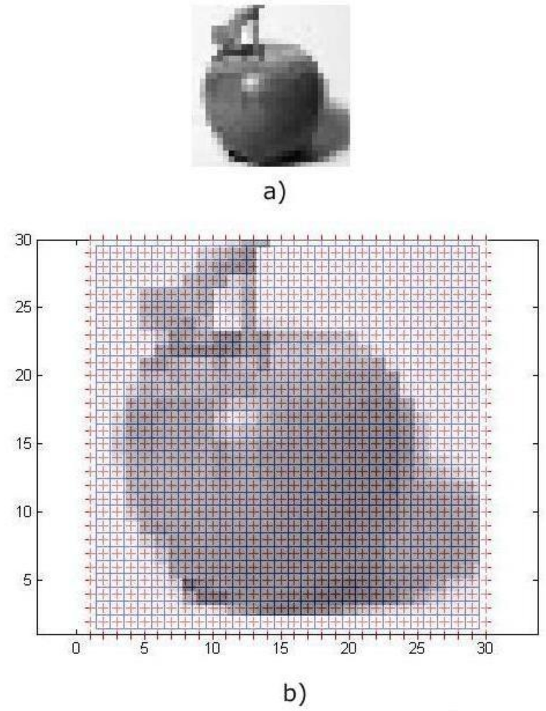

Figure 7: (a)Input Image. (b)Base graph converted from input image.

is encoded as combinatorial map (Fig 7). This graph is the base level of the irregular graph pyramid. The red crosses represent the pixels in the input image, which are the faces in the base graph. Faces are bounded by the blue lines presenting the edges in the graph, and the intersections of the blue lines present the vertices in the graph which are linked by the edges. When the edges are contracted or removed during the building process of irregular graph pyramid, the faces would get merged. So in this pyramid, faces in the graphs represents image regions and the edges of graphs represent the boundaries of image regions.

Each level in the irregular graph pyramid is a graph, while the vertices store their geometric coordinates from the surviving children's position in the original image. With the geometric coordinates of the vertices, we convert the graph in each level into the image of original size for the visualization purpose, see the results in Fig 8. Fig 8 shows the visualization of irregular graph pyramid built on this image. The contraction kernel we select is Maximum Independent Directed Edges Set [14]

Compared to the results from Gaussian Pyramid, the irregular graph pyramid preserves the structure of the object segments, while Guassian Pyramid simply blurred the object by applying low pass filter. The high frequency information, such as the shadow of the stick of the apple survives until higher levels in the irregular graph pyramid. However in Gaussian Pyramid, high frequency information is lost due to low pass filtering building process. In the low resolution (high level) of Gaussian Pyramid, it is nearly impossible to define the shape of the target object while the shape is still preserved in irregular graph pyramid.

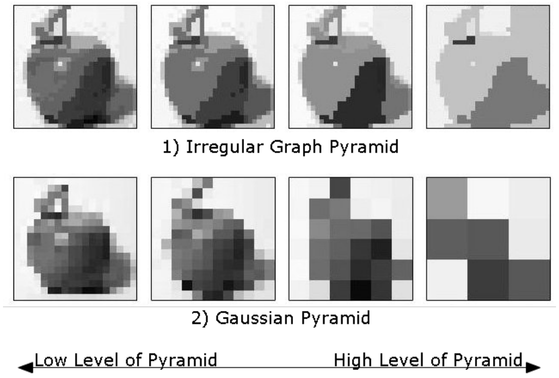

Figure 8: (1)Irregular Graph Pyramid (2) Gaussian Pyramid

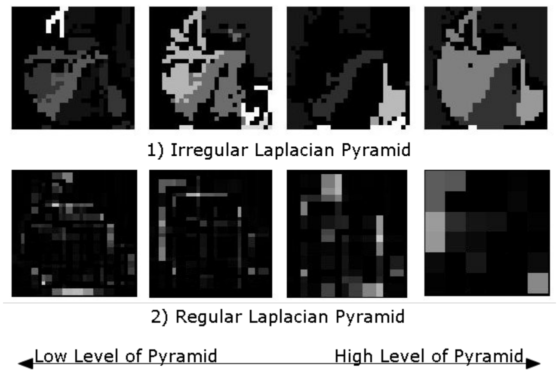

Figure 9: (1) Irregular Laplacian Pyramid. (2) Regular Laplacian Pyramid .

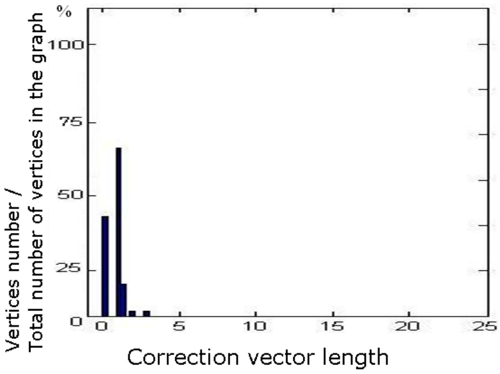

Figure 10: Level 1.

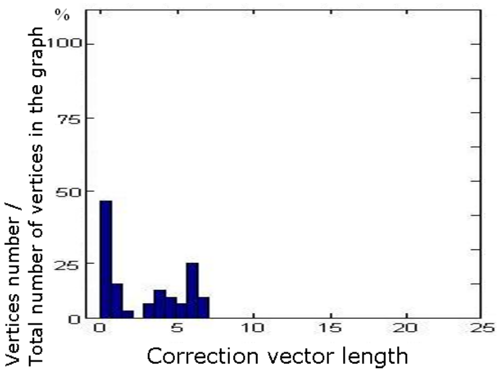

Figure 11: Level 2.

As shown in Fig 9, the results of the irregular Laplacian Pyramid obtain the same advantage as irregular graph pyramid, the advantage of adapting image data into structure, keeping structural information in the pyramid and preserving the topology information.

We may view the regular Laplacian Pyramid as a set of bandpass filtered copies of the original image. Irregular Laplacian Pyramid shows the bandpass filter effect regarding on the length of correction vectors of vertices. As presented in Fig 10, 11 and 12, the length of correction vectors increases as the level of the pyramid increases.

For the graph L_i of level i in the irregular Laplacian Pyramid, N_i refers to the total number of vertices in L_i.

The histograms of level i (Fig 10, Fig 11, Fig 12)show the distribution of vertices with various correction vector length, where horizontal axis presents the length of correction vectors while the vertical axis presents the percentage of vertices with certain length of correction vector respect to the total number of vertices N_i. The histograms show the band pass characteristic regarding to the length of correction vector, which is similar to regular Laplacian Pyramid, as regular Laplacian Pyramid is also band pass filter regarding to the frequency of the image.

In the first level of the pyramid, most of the vertices have correction vector with length 1. Those vertices are the adjacent neighborhood of the surviving vertices, with distance of 1 pixel. The vertices of length 0 are surviving vertices. All the correction vectors length fall in the range between 0 to 5.

In the second level of the pyramid, the overall length of correction vectors increases compared to level 1, with the maximum value exceeds 5.

In the third level of the pyramid, the overall length of correction vectors increase considerably, with the maximum value exceeds 20, because of large regions get merged.

6 Conclusion

This paper presents a novel image representation, irregular Laplacian graph pyramid. It integrates the principles of regular Laplacian Pyramid with the main advantages of irregu-

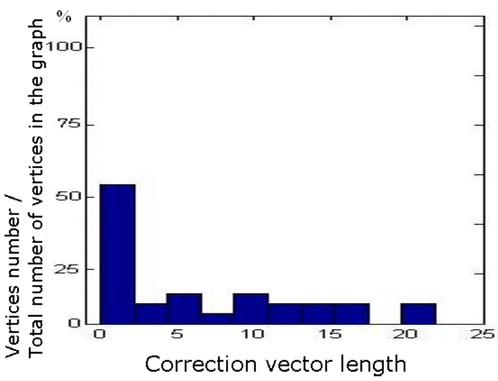

Figure 12: Level 3.

lar graph pyramid. In future work, irregular Laplacian graph pyramid can be applied in image compression and other computer vision problems. The irregular Laplacian Pyramid preserves topological information of target objects. Therefore, this representation may be applied in motion detection, as in which one needs to find the objects movements in the scene.

Acknowledgement

Supported by the Austrian Science Fund under grant P18716-N13

References

[1] P.J. Burt, E.H. Adelson, The Laplacian Pyramid as a Compact Image Code, IEEE Trans. on Communications, pp. 532–540, April 1983. http://citeseer.ist.psu.edu/burt83laplacian.html

[2] K. Knowlton, Progressive Transmission of Gray-Scale and Binary Pictures by Simple, Efficient, and Lossless Encoding Schemes, Proc. of the IEEE, 68(7), 1980, 885-896.

[3] P. J. Burt, Fast filter transforms for image processing, Computer Graphics, Image Processing, vol . 6, pp. 20-51, 1981.

[4] S. B. Lopez Marmol, N. M. Artner, M. Iglesias, W. Kropatsch, M. Clabian, and W. Burger. Improving Tracking Using Structure. In proceedings of Computer Vision Winter Workshop (CVWW 2008). Pages 69 - 76. 4. - 6. February 2008. Moravske Toplice, Slovenia.

[5] J. K. Lee, J. H. Oh, and S. Hwang. Clustering of video objects by graph matching. IEEE International Conference on Multimedia and Expo, pages 394-397, July 2005.

[6] Ana Beatriz V. Graciano, Roberto Marcondes Cesar Junior, Isabelle Bloch: Graph-based Object Tracking Using Structural Pattern Recognition. SIBGRAPI 2007: 179-186

[7] Brun, L. and Kropatsch, W.G. (1999a). Dual Contraction of Combinatorial Maps. Technical Report PRIP-TR-54 Institute f. Computer Aided Automation 183/2, Pattern Recognition and Image Processing Group, TU Wien, Austria.

[8] R. Cori, Un code pour les graphes planaires et ses applications. in: Asterisque, vol. 27. Soc. Math. de France, Paris, France, 1975.

[9] Axel Pinz, Horst Bischof, Walter G. Kropatsch, Gerald Schweighofer, Yll Haxhimusa, Andreas Opelt, Adrian Ion. Representations for Cognitive Vision: A Review of Appearance-Based, Spatio-Temporal, and Graph-Based Approaches,ELCVIA(7), No. 2, 2008.

[10] Guillaume Damiand, Patrick Resch. Split-and-merge algorithms defined on topological maps for 3D image segmentation. Graphical Models 65(1-3): 149-167 (2003)

[11] L. Vacchetti, V. Lepetit, and P. Fua. Stable real-time 3D tracking using online and offline information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(10):1385-1391, 2004.

[12] J. Carranza, C. Theobalt, M. Magnor, and H.-P. Seidel. Freeviewpoint video of human actors. ACM Transaction on Computer Graphics, 22(3), July 2003.

[13] T. Kanade, P. Narayanan, and P. W. Rander. Virtualized reality: Concepts and early results. In IEEE Workshop on the Representation of Visual Scenes, pages 69-76, June 1995.

[14] Yll Haxhimusa, Roland Glantz, Walter G. Kropatsch: Constructing Stochastic Pyramids by MIDES - Maximal Independent Directed Edge Set. GbRPR 2003: 24-34

[15] B. C. R. R. M. L. Matusik, W. and S. Gortler. ImageÍCbased visual hulls. In Proceedings of ACM SIGGRAPH, 2000.

[16] S. Wąğurmlin, E. Lamboray, and M. Gross. 3D video fragments: Dynamic point samples for real-time free-viewpoint video. Computers and Graphics, 28(1):3-14, 2004.