



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Pattern Recognition Letters 26 (2005) 319–337

Pattern Recognition  
Letters

[www.elsevier.com/locate/patrec](http://www.elsevier.com/locate/patrec)

## Vision pyramids that do not grow too high

Walter G. Kropatsch<sup>a,\*</sup>, Yll Haxhimusa<sup>a</sup>, Zygmunt Pizlo<sup>b</sup>, Georg Langs<sup>a,c</sup>

<sup>a</sup> *Institute of Computer Aided Automation, Pattern Recognition and Image Processing Group 183/2,  
Vienna University of Technology, Favoritenstrasse 9, A-1040 Vienna, Austria*

<sup>b</sup> *Department of Psychological Sciences, School of Electrical and Computer Engineering, Purdue University, West Lafayette,  
IN 47907-1364, USA*

<sup>c</sup> *Institute for Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16 2.OG, A-8010 Graz, Austria*

Received 15 July 2004; received in revised form 19 October 2004

Available online 18 December 2004

### Abstract

In irregular pyramids, their vertical structure is not determined beforehand as in regular pyramids. We present three methods, all based on maximal independent sets from graph theory, with the aim to simulate the major sampling properties of the regular counterparts: good coverage of the higher resolution level, not too large sampling gaps and, most importantly, the resulting height, e.g. the number of levels to reach the apex. We show both theoretically and experimentally that the number of vertices can be reduced by a factor of 2.0 at each level. The plausibility of  $\log$  (diameter) pyramids is supported by psychological and psychophysical considerations. Their technical relevance is demonstrated by enhancing appearance-based object recognition. An irregular pyramid hypothesis generation for robust PCA through top-down attention mechanisms achieves higher speed and quality than regular pyramids and non-pyramidal approaches.

© 2004 Elsevier B.V. All rights reserved.

*Keywords:* Image pyramid; Logarithmic complexity; Graph pyramid; Maximal independent set; Robust PCA

### In memoriam

In 1979 the first author met Prof. Azriel Rosenfeld at the international workshop in Maratea,

Italy. The workshop on ‘*Map data processing*’ was a wonderful experience with many opportunities to meet all the famous scientists in the field in the very relaxed atmosphere of this Calabrian village. After Azriel’s talk I dared to ask him a question which was promptly answered, too promptly for me. So I asked him to kindly repeat the answer which he did adding the spelling of the cited authors—it took him not much longer than the first time. . .

\* Corresponding author. Tel.: +43 158801x18350; fax: +43 158801x18392.

*E-mail addresses:* [krw@prip.tuwien.ac.at](mailto:krw@prip.tuwien.ac.at) (W.G. Kropatsch), [yll@prip.tuwien.ac.at](mailto:yll@prip.tuwien.ac.at) (Y. Haxhimusa), [pizlo@psych.purdue.edu](mailto:pizlo@psych.purdue.edu) (Z. Pizlo), [langs@prip.tuwien.ac.at](mailto:langs@prip.tuwien.ac.at) (G. Langs).

In 1984 I got a phone call in my office in Graz, it was Azriel on the phone: “Could you come to my lab?” I do not remember how long my answer was delayed and how I managed to explain that I first had to organize the symposium of German Association for Pattern Recognition (DAGM’84) in Graz. A few weeks after the successful end of DAGM’84 I was in Maryland for a year which I would never want to have missed.

In 1996 we were responsible for the ICPR in Vienna. Of course I asked Azriel to be our honorary chair which he accepted. He participated in a commemorative session organized by Rama Chelapa and held during the banquet in the city hall of Vienna.

A few years ago Azriel established another mutual contact: with Zygmunt Pizlo. Zygmunt approached pyramids from psychophysics and we were glad to find support for our research. The thanks I sent to Azriel came irrecoverably too late. That is why we chose to dedicate our first co-authored paper to the memory of Azriel Rosenfeld.

(Walter G. Kropatsch)

## 1. Introduction

In a regular image pyramid, the number of pixels at any level  $k$  is  $\lambda$  times higher than the number of pixels at the next reduced level  $k + 1$ . The so called reduction factor  $\lambda$  is greater than one and it is the same for all levels  $k$ . If  $s$  denotes the number of pixels in an image  $I$ , the number of new levels on top of  $I$  amounts to  $\log_\lambda(s)$ . Thus, the

regular image pyramid may be an efficient structure to access image objects in a top–down process (see Fig. 1a).

However, regular image pyramids are confined to globally defined sampling grids and lack shift invariance (Bister et al., 1990). In (Montanvert et al., 1991; Jolion and Montanvert, 1992) it was shown how these drawbacks can be avoided by irregular image pyramids, the so called *adaptive pyramids*. Irregular pyramids can perform most of the operations for which their regular counterparts are employed (Rosenfeld, 1987). Each level represents a partition of the pixel set into cells, i.e. *connected subsets of pixels*. The construction of an irregular image pyramid is iteratively local (Meer, 1989; Jolion, 2003):

- the cells have no information about their global position,
- the cells are connected only to (direct) neighbors,
- the cells cannot distinguish the spatial positions of the neighbors.

Although adaptive pyramids overcome the drawbacks of their regular ancestors and although they grow to a reasonable height as long as the base is small, they grow higher than the base diameter with a larger input size because the progressive deviation from the regular base favors configurations that slow down the contraction process. As a consequence of the greater height the efficiency of pyramids degrades. We show that this problem can be resolved by a new selection mechanism which guarantees logarithmic heights.

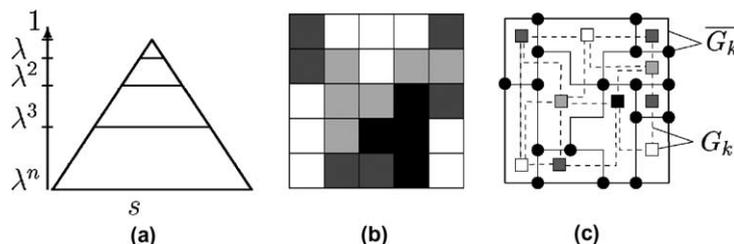


Fig. 1. (a) Pyramid concept, reduction factor and (b) partition of pixel set into cells image. (c) Representation of the cells and their neighborhood relations by a dual pair  $(G_k, \overline{G}_k)$  of plane graphs.

We replace the selection method for contraction kernels proposed in (Meer, 1989) by two new iteratively local methods: MIES and MIDES. The method MIES (Haxhimusa et al., 2002) guarantees a reduction factor of at least 2.0 and the method MIDES (Haxhimusa et al., 2003) for the cases where there are constraints on the direction of contraction. Experiments with selection methods show that the method in (Meer, 1989) does not lead to logarithmic tapering graph pyramids, as opposed to our methods, i.e. the reduction factors of graph pyramids built by this method can get arbitrarily close to 1.0, which means that the *height* of the pyramid can be very high. This is not in contradiction to the reported average reduction factor of more than 4 because Meer uses 8-connectivity and can not exclude particular cases where pyramids grow too high. This problem is encountered also in (Montanvert et al., 1991). Not only stochastic decimation (Meer, 1989) but also connected component analysis (Kropatsch and Macho, 1995) gain from the new methods. The method in Section 6 turned out to produce logarithmic tapering graph pyramids also in case of monotonic dual graph contraction (Glantz and Kropatsch, 2000).

Irregular pyramids can be used to enhance appearance based object recognition. Their structure gives the opportunity to apply top-down processing on data and introduce additional a priori knowledge similar to visual attention guidance. Thereby efficiency of succeeding algorithms can be improved considerably.

The plan of this paper is as follows: We first give some strong motivation for the plausibility of a pyramidal architecture in biological vision (Section 2). These hierarchies must be ‘shallow’ to be efficient in both bottom-up information aggregation and in top-down processes for verification and focus of attention. In Section 4 we recall the main selection algorithm used in the stochastic pyramid construction and demonstrate in Section 7.1 that graph pyramids from maximal independent vertex sets may have a very poor reduction factor. Moreover, experiments show that small reduction factors are likely, especially when the images are large. We propose two modifications. MIES in Section 5 guarantees a reduc-

tion factor of at least 2.0, proven theoretically, but is applicable only if the edges may be contracted in both directions. The modification proposed in Section 6 (MIDES) also works in case of constraints on the directions of contraction. This modification yields the highest reduction factors in the case of stochastic graph pyramids in all our tests, as our experiments in Section 7 confirm. In the Eigen-pyramid approach the irregular pyramid has been used efficiently to improve the speed and the quality of appearance-based recognition (Section 8).

## 2. Pyramids in human vision

It is now quite generally accepted that the human visual system has a pyramid architecture and that the visual mechanisms can be adequately modeled by pyramid algorithms. Specifically, neurophysiological and neuroanatomical data indicates that the visual systems of cats, monkeys and human beings are hierarchical, with neurons on lower layers having smaller receptive fields and neurons on higher layers having larger receptive fields (Zeki, 1993). Experiments characterizing receptive fields of neurons in the visual system started with Kuffler’s seminal paper published in 1953 (Kuffler, 1953), and by the end of the 1960s a number of facts concerning the architecture of the visual system had been established. In parallel to physiological studies, psychophysicists were trying to formulate computational models of the perceptual mechanisms. One of the first such models was proposed in 1968 by Campbell and Robson (1968), who conjectured that the human visual system performs Fourier analysis of spatial relations on the retina. This conjecture was equivalent to the assumption that the visual system consists of a set of spatial frequency channels, whose output is integrated by means of linear operations. This model did capture the multiresolution nature of the visual system, and has been accepted as a plausible theory for at least a quarter of century. It is somewhat surprising, however, that this model received so much attention considering its fundamental limitations such as lack of spatial localization and the linearity of the operations.

Shortly after Campbell and Robson published, Aziel Rosenfeld published his first paper on pyramids (Rosenfeld and Thorston, 1971). Rosenfeld realized that effective image processing algorithms must involve operators on several levels of scale and resolution. Rosenfeld was certainly aware of the Fourier analysis model of vision. However he, unlike others, recognized very early the importance of spatial localization in vision and the critical role of non-linear operations (Rosenfeld, 1970). For reasons that are not entirely clear, pyramid algorithms, despite their physiological and psychophysical plausibility, have not received much attention by modelers of human vision. They were first adopted as a standard tool by those psychophysicists whose interests were close to image processing, rather than image understanding (the reader may verify this claim by consulting the human vision journals such as *Vision Research* and *Perception and Psychophysics*).

It is important to emphasize that by “pyramid algorithms” we mean any computational tool that performs image analysis based on multiple representations of the image forming a hierarchy with different scales and resolution, and in which the height (number) of a given level is a logarithmic function of the scale (and resolution) of the operators. Multiresolution pyramids form a subset of the general class of exponential pyramid algorithms. The multiresolution pyramids perform only linear operations like average, difference, etc. Pyramid algorithms, which incorporate a wider class of operators, are adequate models for the Gestalt rules of perceptual organization such as proximity, good continuation, common fate (Pizlo et al., 1997; Pizlo, 2001). They also provide an adequate model of Weber’s law and the speed-accuracy trade-off in size perception, as well as of the phenomenon of mental size transformation (Pizlo et al., 1995). In the case of size processing, modeling visual processes involves both bottom-up (fine to coarse) and top-down (coarse to fine) analysis. The top-down processing seems also critical in solving the image segmentation problem, which is a difficult inverse problem (Bouman and Liu, 1991). This problem has received much attention in psychological literature, and is known as

figure-ground segregation phenomenon (Koffka, 1935).

The human visual system takes advantage of the ability to distinguish between highly valuable and less relevant regions in the field of view. Thereby it improves its performance considerably. Stark and Privitera (1997) and Chernyak and Stark (2001) describe two strategies to obtain and apply information about the importance of different regions of an image when simulating the human visual system. Bottom-up methods retrieve their features only from the present input image (Privitera and Stark, 2000). Top-down methods are driven by knowledge which is available before getting the input. Experiments (Stark and Privitera, 1997) have shown that human vision and particularly the scan paths of the eyes, composed of saccades, are not only dependent on the input image, but largely on previous knowledge i.e. top-down expectations.

More recently, pyramid algorithms have been used to model the mental mechanisms involved in solving the visual version of the Traveling Salesman Problem (Graham et al., 2000), as well as other types of visual problems (Pizlo and Li, 2003, 2004). Humans seem to represent states of a problem by clusters (recursively) and determine the sequence of transformations from the start to the goal state by a top-down sequence of approximations. This approach leads to algorithms whose computational complexity is as low as that of the mental processes (i.e. linear), and which produce solution paths that are close to optimal. It follows that pyramid models may provide the first plausible explanation of the phenomenon of the directedness of thought and reasoning (Humphrey, 1948).

### 3. Graph pyramid

On the base level (level 0) of an irregular image pyramid the cells represent single pixels and the neighborhood of the cells is defined by the 4(8)-connectivity of the pixels. A cell on level  $k + 1$  (parent) is a union of some neighboring cells on level  $k$  (children). This union is controlled by so called *contraction kernels* (CK) (Kropatsch, 1995). Every parent computes its values independently of other cells on the same level. This leads

to the property that an image pyramid is built in  $O[\log(\text{image\_diameter})]$  time. For more information on the subject, see the book of Jolion and Rosenfeld (1994) and of Rosenfeld (1984).

Neighborhoods on level  $k + 1$  are derived from neighborhoods on level  $k$ . Two cells  $c_1$  and  $c_2$ , at level  $k + 1$ , are neighbors if there exist pixels  $p_1$  in  $c_1$  and  $p_2$  in  $c_2$ , at level  $k$  such that  $p_1$  and  $p_2$  are 4(8)-neighbors. We assume that on each level  $k + 1$  ( $k \geq 0$ ) there exists at least one cell not contained in level  $k$ . In particular, there exists a highest level  $h$ . Furthermore, we restrict ourselves to irregular pyramids with an apex, i.e. level  $h$  contains only one cell.

A level consists of dual pair  $(G_k, \overline{G}_k)$  of plane graphs  $G_k$  and  $\overline{G}_k$ , Fig. 1c. The planarity of graphs restricts us to using only the 4-connectivity of the pixels. The vertices of  $G_k$  represent the cells on level  $k$  and the edges of  $G_k$  represent the neighborhood relations of the cells on level  $k$ , depicted with square vertices and dashed edges in Fig. 1c. The edges of  $\overline{G}_k$  represent the borders of the cells on level  $k$ , solid lines in Fig. 1c, possibly including so called pseudo edges needed to represent neighborhood relations to a cell completely enclosed by another cell. Finally, the vertices of  $\overline{G}_k$ , circles in Fig. 1c, represent meeting points of boundary segments of  $\overline{G}_k$ , solid lines in Fig. 1c. The sequence  $(G_k, \overline{G}_k)$ ,  $0 \leq k \leq h$  is called (dual) graph pyramid.

**Definition 1.** The *reduction factor*  $\lambda$  is the ratio of the number of vertices of graphs  $G_{k+1}(V_{k+1}, E_{k+1})$  and  $G_k(V_k, E_k)$ :  $|V_{k+1}| \leq \frac{|V_k|}{\lambda}$ .

The aim of using graphs is to combine the advantage of regular pyramids (logarithmic tapering) with the advantages of irregular pyramids (their purely local construction, universal segmentation, topology preservation, preservation of face degree, etc.). The aim is reached by replacing the selection method for contraction kernels proposed in (Meer, 1989) by two new iteratively local methods: MIES and MIDES.

#### 4. Maximal independent vertex set (MIS)

In the following the iterated local construction of the (stochastic) irregular image pyramid in

(Meer, 1989) is described in the language of graph pyramids. The main idea is to first calculate a so called *maximal independent vertex set*<sup>1</sup> (Christofides, 1975). Let the vertex set and edge set of  $G$  be denoted by  $V$  and  $E$ , respectively. The incidence relation of  $V$ , denoted by  $\iota(\cdot)$  maps each edge from  $E$  to its set of end vertices.

The neighborhood  $\Gamma(v)$  of a vertex  $v \in V$  is defined by

$$\Gamma(v) = \{\{v\} \cup \{w\} \in V \mid \exists e \in E \text{ such that } v, w \in \iota(e)\}.$$

**Definition 2.** A subset  $W$  of  $V$  is called a *maximal independent vertex set* if:

1.  $w_1 \notin \Gamma(w_2)$  for all  $w_1, w_2 \in W$ ,
2. for all  $v \in V$  there exists a vertex  $w \in W$  such that  $v \in \Gamma(w)$ .

The 1st condition requires that any two surviving vertices cannot be in the neighborhood of each other (black vertices in Fig. 2a and b). The 2nd condition says that every non-surviving vertex has in its neighborhood a surviving vertex (white vertices in Fig. 2a and b). A maximal independent vertex set  $W$  is not necessarily maximum, as there may be another set  $W'$  that contains more vertices than  $W$  (Fig. 2b).

##### 4.1. Algorithm MIS

The maximal independent vertex set (MIS) problem was solved using a heuristic in (Meer, 1989). The number of iterations to complete MIS converges in most of the cases very fast, so called *iterations for correction* (Meer, 1989). MIS (Meer, 1989; Jolion, 2003) may be generated as shown in Algorithm 1. We assume that no two random numbers are equal.

<sup>1</sup> Also called maximal stable set; we distinguish maximal from maximum independent set, whose construction is NP-complete (Thulasiraman and Swamy, 1992).

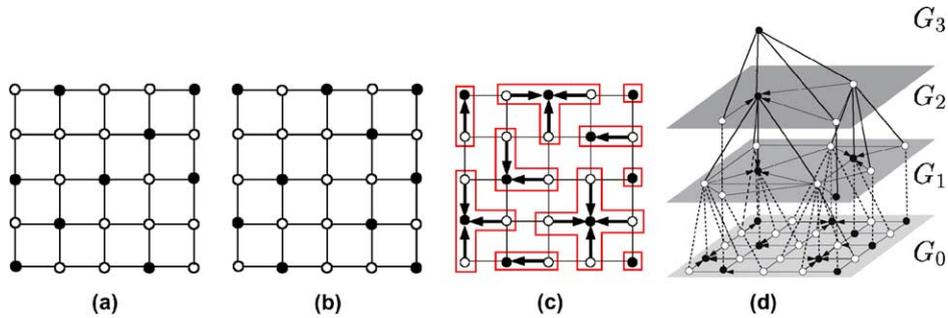


Fig. 2. (a) A maximal independent vertex set  $W$  with 9 vertices, (b) a maximal independent vertex set  $W'$  with 10 vertices, (c) contraction kernels corresponding to  $W'$ , and (d) irregular graph pyramid built using the contraction kernels from (c).

**Definition 3.** A *contraction kernel* is a rooted tree. The depth of a kernel is the longest distance of a leaf from the root, the diameter is the longest path through the tree.

With this definition we can redefine the reduction factor as follows:

**Proposition 4.** The reduction factor can be determined from the sizes of vertices and the number of contraction kernels  $|\mathcal{C}| : \lambda \geq \frac{|V|}{|\mathcal{C}|}$ .

**Proof.** The roots of each contraction kernel from  $\mathcal{C}$  are the survivors in the next level of the pyramid. Hence the number of vertices in the next level is the same as the number of contraction kernels, i.e.  $|V_{k+1}| = |\mathcal{C}|$ .  $\square$

#### Algorithm 1 (Algorithm MIS)

**Input:** Graph  $G(V, E)$ .

1. Mark every element of  $V_k$  as *candidate*.
2. **While** there are candidates **do**
3. Assign random numbers to the candidates of  $V_k$ .
4. Determine the candidates whose random numbers are greater than the random numbers of all neighboring candidates and mark them as *member* (of the maximal independent set) and as *non-candidate*.
5. Mark every neighbor of every new member as *non-candidate*.

6. In each neighborhood of a vertex that is not a member there will now be a member. Let each non-member choose the edge to be contracted to its neighboring member, say the one with the maximal random number.

**Output:** Set  $\mathcal{C}$  of contraction kernels based on MIS.

The assignment of the non-members to their members determine a collection of *contraction kernels*  $\mathcal{C}$ : each non-member is contracted towards its member and all contractions can be done in a single parallel step. In Fig. 2c the contractions are indicated by arrows. A graph pyramid from maximal independent vertex sets is shown in Fig. 2d, where  $G_i$  are graphs in levels  $i = 0, \dots, 3$ . This can be done by the dual graph contraction algorithm (Kropatsch, 1995). Note that we remove parallel edges and self-loops that emerge from the contractions if they are not needed to encode inclusion of regions by other regions. In the example of Fig. 2d we do not need loops nor parallel edges. Trivial contraction kernels occur very often in MIS as can be seen in Fig. 2c: isolated black vertices. In Section 5 we introduce a method that finds contraction kernels that are non-trivial.

#### 5. Maximal independent edge set (MIES)

In the following we aim at a collection  $\mathcal{C}$  of contraction kernels in a plane graph  $G$  such that:

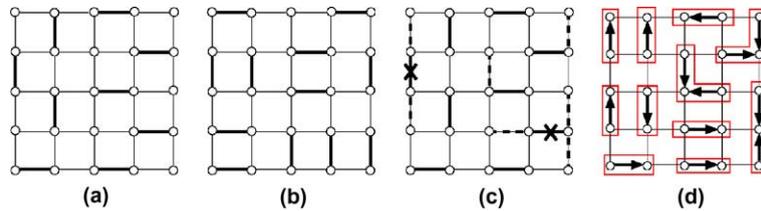


Fig. 3. (a) A maximal matching  $M$  with 9 edges, (b) a matching  $M'$  with 12 edges, (c) enlarged matching  $M^+$  from  $M$  by connecting formerly isolated vertices to the maximal matching  $M$ , after breaking up trees of diameter three into trees of depth one, and (d) the frames indicate a corresponding collection of contraction kernels.

- each vertex of  $G$  is contained in exactly one kernel of  $\mathcal{C}$ , and
- each kernel in  $\mathcal{C}$  contains at least two vertices.

**Definition 5.** A set  $M$  of independent edges<sup>2</sup> in a graph  $G = (V, E)$  is a *maximal matching* if every vertex in  $U \subset V$  is incident with an edge on  $M$ , and  $M$  cannot be enlarged by an edge without losing independence. The vertices in  $U$  are then called *matched vertices* (by  $M$ ); all other vertices are called *unmatched vertices*.

A maximal matching  $M$  is not necessarily maximum: there may be a matching  $M'$  (Fig. 3b) that contains more edges than  $M$  (Fig. 3a). We assume that  $G$  is connected and planar, and that the data do not impose any constraint on the selection, for e.g. in large homogeneous regions. Clearly, the contraction of all kernels in  $\mathcal{C}$  will reduce the number of vertices to half or less.

5.1. Algorithm MIES

We start with independent *edge sets* or *matchings*, i.e. edge sets in which no pair of edges has a common end vertex. The *maximal independent edge set* (MIES),  $\mathcal{C}$  is constructed in three steps (Algorithm 2).

A maximal matching of  $G$  is equivalent to a maximal independent vertex set on the edge-graph of  $G$  (Diestel, 1997; Christofides, 1975). Thus, a

maximal matching may be determined by the iterated process used in MIS algorithm (Section 4). Note that  $M$  is only required to be maximal, i.e. the edge set  $M$  cannot be enlarged by another edge from  $E \setminus M$  without losing independence.

**Algorithm 2** (Algorithm MIES)

**Input:** Graph  $G(V, E)$ .

1. Find a maximal matching  $M$  of edges from  $G$ .
2.  $M$  is enlarged to a set  $M^+$  that spans a subgraph of  $G$ .
3.  $M^+$  is reduced to a subset defining  $\mathcal{C}$ .

**Output:** Set  $\mathcal{C}$  of contraction kernels based on MIES.

The collection of contraction kernels defined by a maximal matching  $M$  may include kernels with a single vertex. Let  $v$  denote such an isolated vertex, isolated from  $M$ , and choose a non-self-loop  $e$  that has  $v$  as an end vertex. Since  $M$  is maximal, the other end vertex  $w \neq v$  of  $e$  belongs to an edge that is contained in the matching  $M$ . Let  $M^+$  denote the set of edges that are in  $M$  or that are chosen to connect isolated vertices to  $M$ , 2nd step of MIES.

The subgraph of  $G$  that is induced by  $M^+$  spans  $G$  and its connected components are trees of diameter one, two or three, Fig. 3c. A tree of diameter three can be separated into two trees of diameter one each by removing the unique edge, both end vertices of which belong to other edges of the tree (indicated by the crosses in Fig. 3c); the 3rd step of MIES. Still, each vertex of  $G$  belongs to a tree (of depth one). Since each vertex of  $G$  is now contained in a non-trivial contraction kernel, we prove the following proposition:

<sup>2</sup> Two edges are independent if they are not incident on the same vertex.

**Proposition 6.** *The number of contraction kernels produced by algorithm MIES is at most  $|V|/2.0$ .*

**Proof.** Let  $G = (V, E)$  be a planar connected graph with  $|V_k|$  vertices. Let  $M$  be a maximal matching of  $G$  and let  $U$  be the set of vertices matched by  $M$  and  $U' = M \setminus U$  be the set of unmatched vertices (isolated vertices). A vertex  $v \in U$  is called matched by  $M$  if it is incident with an edge in  $M$ . Since the matching is maximal, no two adjacent vertices  $v, w$  may be unmatched, i.e. in the neighborhood of the unmatched vertex there is at least one matched vertex, if the vertex is not already isolated in  $G$ . This means that in the neighborhood of an unmatched vertex  $w \in U'$  there is at least one edge connecting  $w$  to a matched vertex  $v \in U$ . If there is more than one connecting edge, we select only one of them.

The edge set  $M$  is enlarged to  $M^+$  by adding all connecting edges. Thus all vertices of  $V$  are incident to edges  $M^+$ . The subgraph of  $G$  that is induced by  $M^+$  spans  $G$  and its connected components are trees of diameter one (the matched edge), two (isolated vertices connected to one endpoint of a matched edge only) or three (isolated vertices connected to both ends of the matched edge). Trees of diameter three are split into trees of diameter one by removing the *unique central edge*  $ue \in U$ . The two endpoints of the matched edge of a tree of diameter three have a degree of at least two (they cannot be leaves). Splitting of the tree removes this matched edge which reduces the degree of both end vertices by one, but they are still greater than zero. Hence this edge deletion does not create a new isolated vertex. This ensures that all vertices  $V_k$  are also incident edges of the new set  $M^{++} = M^+ \setminus \{ue\}$  and there are no isolated vertices left.

We conclude that the subgraph of  $G$  that is induced by  $M^{++}$  spans  $G$  and its connected components are trees of diameter one or two i.e. with more than one vertex. All the components are trees and have diameter one or two and the number of contraction kernels  $\mathcal{C}$  is  $|V|/2.0$ . If there is at least one tree of diameter two then  $|\mathcal{C}| < |V|/2.0$ . We proved that in the general case  $|\mathcal{C}| \leq |V|/2.0$ .  $\square$

The surviving vertices can be in the neighborhood of each other, which means that the 1st con-

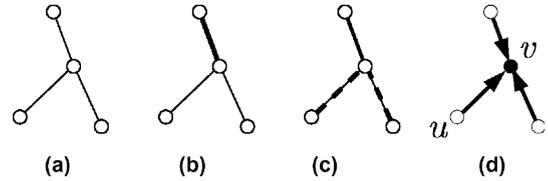


Fig. 4. Restriction in choosing the surviving vertex. (a) Input graph, (b)  $M$ : maximal matching. (c)  $M^+$ : the matching from (a) enlarged by connecting formerly isolated vertices to the maximal matching. (d) The only possible CK.

dition of Definition 2 is relaxed; this is proposed in (Kropatsch and Montanvert, 1991). Non-surviving vertices are in the neighborhood of at least one surviving vertex, fulfilling the 2nd condition of Definition 2.

Note that for the input graph in Fig. 4a, the surviving vertex cannot be chosen arbitrarily. In this case the algorithm chooses the only possible vertex  $v$  for the root of the tree i.e. as a survivor and there is only one possible contraction kernel, implying also the direction of contraction. For other cases the survivor can be chosen arbitrarily. The arrows in Fig. 3d indicate possible directions of contractions, i.e. possible choice of the survivors. But some applications (Burge and Kropatsch, 1999; Kropatsch and Burge, 1998; Glantz et al., 1999; Kammerer and Glantz, 2001; Glantz and Kropatsch, 2000) could restrict the way of choosing the surviving vertices, i.e. choosing vertex  $u$  as a survivor, or could constrain the direction of contraction. This is why the proposed method cannot be extended to applications in which there are a priori constraints on the directions of the contractions. However, the proposed method works for the stochastic case (no preconditions on the direction of edges to be contracted) and for connected component analysis (Kropatsch and Macho, 1995), where the attributes of the end vertices are required to be identical.

## 6. Maximal independent directed edge set (MIDES)

In many graph pyramid applications such as line image analysis (Burge and Kropatsch, 1999; Kropatsch and Burge, 1998) and the description of image structure (Glantz et al., 1999; Kammerer

and Glantz, 2001; Glantz and Kropatsch, 2000) a directed edge  $e$  with source  $u$  and target  $v \neq u$  must be contracted from  $u$  to  $v$ , only if the attributes of  $e$ ,  $u$ , and  $v$  fulfill a certain condition. In particular, the condition depends on  $u$  being the source and  $v$  being the target, making the direction of contraction an important issue (Fig. 5a). In line drawings the end point of a line or an intersection must be preserved for geometric accuracy reasons. The edges that fulfill the condition are called *preselected* edges.

From now on the plane graphs in the pyramid have (bi)directed edges. Typically, the edges in the base level of the pyramid form pairs of reverse edges, i.e. for each edge  $e$  with source  $u$  and target  $v$  there exists an edge  $e'$  with source  $v$  and target  $u$ . Undirected graphs can be converted into directed graphs by substituting the edges into pairs of reverse edges. However, the set of preselected edges may contain  $e$  without containing  $e'$ . The goal is to build contraction kernels with a “high” reduction factor from the set of preselected edges. The reduction will always be determined according to the directed graph induced by the preselected edges. For example, if the number of vertices in the induced subgraph is reduced to half, the reduction factor will be 2.0. From the example in Fig. 5b it is clear that, in general, the reduction factor can be arbitrarily close to 1.0. To perform the contractions in parallel, we need a vertex disjoint union of contraction kernels. The plan is to define such a

union in terms of independent directed edges, where “independent” means that no pair of directed edges belongs to the same neighborhood  $N(e)$ . Then, dealing with edges instead of vertices, we may find the contraction kernels as in MIS.

**Definition 7.** Let  $e = (u, v)$  be a directed edge of  $G$  and  $u \neq v$ . Then the directed neighborhood  $N(e)$  is given by all directed edges with the same source  $u$ , targeting the source  $u$  or emanating from target  $v$ :

$$N(e) = N((u, v)) = \{(u, v') \in E\} \cup \{(u', u) \in E\} \cup \{(v, u') \in E\}.$$

The neighborhood  $N(e)$  of  $e$  is given by all edges which point towards the source of  $e$ , all edges with the same source  $u$  and all the edges the source of which is the target of  $e$ . Note that edges pointing towards the target of  $e$  are not part of the directed neighborhood. Seen from a directed edge  $e$  with source  $u$  and target  $v \neq u$  that one wants to contract (from  $u$  to  $v$ ), no edge  $e' \neq e$  with end vertex (source or target) equal to  $u$  or source equal to  $v$  may be contracted. Fig. 5e depicts  $N(e)$  in case of  $u$  and  $v$  both having 4 neighbors.

Note that the direction of edges uniquely determines which vertex survives on the next level of the pyramid, i.e. determines the contraction kernel (decimation parameter Kropatsch, 1995). A set  $\mathcal{C}$  of directed edges forms such a collection of contraction kernels if and only if  $\mathcal{C}$  contains none of

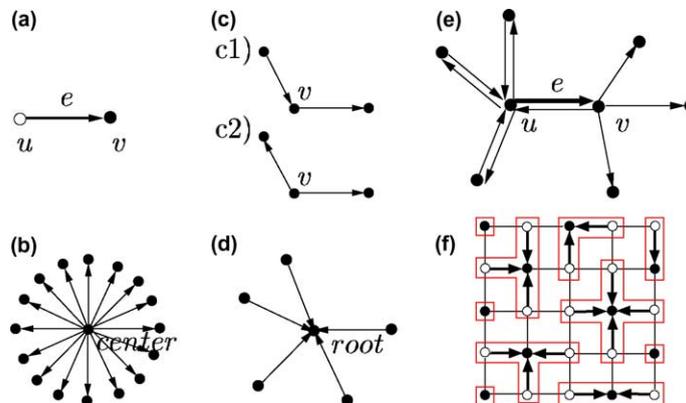


Fig. 5. (a) The direction of contraction. (b) The reduction factor of a star with  $n$  edges pointing away from the center is  $(n + 1)/n$ . (c) Pairs of directed edges forbidden in a contraction kernel. (d) A legal configuration of directed edges. (e) The neighborhood  $N(e)$ . (f) CKs built by maximal independent edge set with respect to  $N(e)$ .

the edge pairs depicted in Fig. 5c. An edge  $e$  to be contracted together with those edges that one may not contract form a directed edge neighborhood  $N(e)$  of  $e$ .

**Definition 8.** Let  $v$  be a vertex of a bi-directed graph  $G = (V, E)$ . We define the out-degree as

$$s(v) := |\{e \in E | v \text{ is source of } e\}|,$$

and the in-degree as

$$t(v) := |\{e \in E | v \text{ is target of } e\}|.$$

In Fig. 5d the number of edges with target in *root* is  $t(\text{root}) = 5$ ; and for the center vertex with  $n$  edges pointing away  $s(\text{center}) = n$ , Fig. 5b.

**Proposition 9.** Let  $D$  denote a subset of directed edges of a bi-directed graph  $G$  and  $G_D$  denote the subgraph induced by  $D$  with out-degrees  $s_D(\cdot)$  and in-degrees  $t_D(\cdot)$ . Then the following statements are equivalent:

- (a)  $s_D(v) < 2 \wedge s_D(v) \cdot t_D(v) = 0, \quad \forall v \in G_D$ .
- (b)  $G_D$  is a vertex disjoint union of contraction kernels.

**Proof.** (i) (a)  $\Rightarrow$  (b): Let  $R := \{r \in G | r \text{ is the target of some } e \in D \wedge s(r) = 0\}$  be the set of roots. Furthermore,  $E_r := \{e \in D | r \text{ is target of } e\}$  is a tree with root  $r \in R$ , hence a contraction kernel  $\mathcal{C}_r$ . It remains to show that any two contraction kernels  $\mathcal{C}_v$  and  $\mathcal{C}_w, u \neq w \forall v, w \in R$ , are vertex disjoint. Assume the opposite, i.e. there exists a vertex  $u$  contained in  $\mathcal{C}_v$  and in  $\mathcal{C}_w$  for some  $v \neq w \in R$ . There are two cases:

- (1)  $u$  is a root and  $u = v = w$  since every contraction kernel has a single root enforcing  $v = w$ ; this contradicts  $v \neq w$ ;
- (2)  $u$  is a leaf hence, there exist edges  $(u, v)$  and  $(u, w) \in D$ , a contradiction to  $s(u) < 2$ .

(ii) (b)  $\Rightarrow$  (a): Let  $T$  be the set of roots of the vertex disjoint contraction kernels and let  $C_t$  denote the unique contraction kernel with root  $t, t \in T$ . Furthermore, let  $v \in G_D$ . Since the  $\mathcal{C}_t$  are vertex disjoint, exactly one of the following holds  $\forall v \in V$ :

- (1)  $v \in T$  is a root and  $s(v) = 0$ .
- (2)  $v \notin T$  is a leaf of a tree  $C_t$  and  $s(v) = 1, t(v) = 0$ .  $\square$

Examples of kernels which do not fulfill the Definition 3 are shown in Fig. 5c1 where  $s(v) = 1$  and  $t(v) = 1$ ; and Fig. 5c2 where  $s(v) = 0$  and  $t(v) = 2$ . From the example in Fig. 5d it is clear that only one edge can be contracted (otherwise one ends up with forbidden contraction kernels), which means in general the vertex reduction factor can get arbitrarily close to 1.0.

Note that, in contrast to MIS, the roots of two contraction kernels may be neighbors. Condition (a) in Proposition 9 is fulfilled if and only if no pair of directed edges from  $D$  belongs to the same  $N(e), e \in D$ . Hence, a maximal vertex disjoint union of contraction kernels may be found via a maximal set of directed edges that are independent with respect to  $N(e)$ . A parallel algorithm to find a maximal independent set with respect to  $N(e)$  is specified in the next section.

### 6.1. Algorithm MIDES

To find a *maximal (independent) set of directed edges (MIDES)* forming vertex disjoint rooted trees of depth zero or one, we proceed analogously to the generation of maximal independent vertex sets (MIS), as explained in Section 4. Let  $E$  denote the set of directed edges in the graph  $G$  of the graph pyramid. We proceed as in Algorithm 3.

#### Algorithm 3 (Algorithm MIDES)

**Input:** Graph  $G(V, E)$ .

1. Mark every directed edge of  $E$  as *candidate*.
2. **while** there are candidates
3. Determine the candidates  $e$  whose random numbers are higher (larger) than the random numbers in  $N(e) \setminus \{e\}$  and mark them as *member* (of a contraction kernel). Also mark every  $e' \in N(e)$  of every new member  $e$  as *non-candidate*.
4. The target of the directed edges are marked as survivors, all the vertices which are the sources of directed edges are marked as non-

survivor and the remaining unmarked vertices are marked to be survivors as well.

**Output:** Set  $\mathcal{C}$  of contraction kernels based on MIDES.

Since the direction of edges uniquely determines the roots of the contraction kernels (the survivors), all the vertices which are the sources of directed edges are marked as non-survivors. The remaining isolated vertices are marked to be survivors. An example of a set of contraction kernels  $\mathcal{C}$  found by MIDES is given in Fig. 5f (the survivors are depicted with black and non-survivors with white).

Isolated vertices  $i$  can occur after MIDES, too. Because all contraction kernels are disjoint and because each edge covers two vertices it is clear that following proposition holds:

**Proposition 10.** *If there are no isolated vertices left after algorithm MIDES the number of contraction kernels is at most  $|V|/2$  (allowing a reduction factor of at least 2.0).*

In order to understand the difference to MIES let us study what happens in the while-loop after the first iteration. After local maxima are marked as members and all neighbors are removed as candidates, all remaining edge candidates are non-maxima.

The particular neighborhood  $N(e)$  used in the (bi-)directed graph creates the edge adjacencies shown in the edge-graph (Fig. 6a). Each directed edge corresponds to a vertex ( $\square$ ) in the edge-graph. All edges in the neighborhood  $N(e)$  create edges in the edge-graph. Two edges pointing to the same vertex are NOT connected in the edge-graph.

Isolated vertices  $i$  appear at the end of the correction phase when the neighborhood of edge-

members cover all edges incident to  $i$ . Furthermore all edges incident to  $i$  are edge-neighbors of a member whose orientation points away from  $i$ . Otherwise  $N(e)$  would allow the reverse edge to be a member. In Fig. 6a edges  $a, 0$  are incident to the isolated vertex  $i$ , both neighbors of edge  $b$ ,  $N(b) = \{0, a, c, d\}$ . If  $c$  would be the member instead of  $b$ , then  $0$  would become a member, too, since  $0 \notin N(c) = \{e, d, b, a\}$  and  $i$  would not be isolated.

From each isolated vertex  $i$  emanate paths with monotonically increasing random numbers leading to a local maximum. We therefore study such a path which is a linear sequence of vertices (Fig. 6a).

Assume  $g$  is the only local maximum, all the other edges have at least one edge-neighbor which has a higher value. In this case  $g$  would become a member and  $e, f$  would be removed as candidates. The next highest edges could be  $c$  or  $d$ . If  $c$  is higher then it would be selected next and  $a, b, d$  removed leaving  $0$  which would be the last edge selected. In this case no isolated vertex would appear. Assume  $d$  is selected instead of  $c$ . Then  $b, c$  would be removed as neighbors,  $a$  would be selected next and  $0$  removed as neighbor of  $a$ . We see that many possibilities lead to solutions without an isolated vertex.

Let us reason backwards: In order for vertex  $i$  to be an isolated vertex, both edges  $0$  and  $a$  must be neighbors of a member. The only choice is  $b!$

Case  $c, d > b$ : If both  $c$  and  $d$  are higher than  $b$  then  $e$  is the last member before  $b$  was chosen. Since  $b$  and  $e$  share the same root vertex they are part of the same contraction kernel covering 3 vertices. Together with the isolated vertex the two kernels cover 4 vertices giving a reduction factor of 2 (see Fig. 6b).

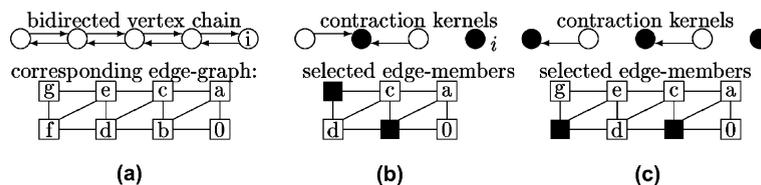


Fig. 6. The edge-graph of a sequence of vertices: (a) labeled edge-graph, (b) case  $c, d > b$ , (c) case  $c < b$ .

Case  $c < b$ : If  $c$  is less than  $b$  then  $f$  could be the last edge-member before  $b$  (see Fig. 6c). We note that the selected edges for contraction point in the same direction.

We can continue the construction until either the orientation of the selected edge changes or we reach the local maximum. In the first case the two reversed edges form a contraction kernel covering three vertices and, hence compensating the isolated vertex. In the second case the search for a larger contraction kernel covering more than just two vertices would continue for all other paths emanating from the same isolated vertex. The only case where the isolated vertex  $i$  is not compensated by a larger contraction kernel consists of a set of edge-members filling the region of the correction phase and ALL pointing away from  $i$  and towards the local maximum. As we will see in our experiments this is extremely unlikely and could, in addition, be compensated by any other larger contraction kernel.

MIDES has another feature different from MIES. It can create star-like kernels due to the special neighborhood chosen. Edges  $(x_1, r)$ ,  $(x_2, r), \dots$  having the same target  $r$  can ALL be local maxima and selected as members in the first iteration of algorithm MIDES since  $N((x_1, r)) \cap N((x_2, r)) = \emptyset$ . Therefore the resulting set of edge-members is not necessarily a matching and may well contain larger kernels with one root and several edges attached to it.

We conjecture that larger kernels occur more frequently than isolated vertices. Each isolated vertex needs only one kernel of more edges to keep the balance for a reduction factor of two. This may explain the surprising experimental results with MIDES.

## 7. Comparing the speed of reduction

Uniformly distributed random numbers, as the simplest choice of independent identically distributed (i.i.d.) random numbers, are assigned as attributes to the vertices (MIS) or edges (MIES, MIDES) in the base level grid graphs. By changing

the seed of the uniformly distributed random generator we generated 1000 graphs, on top of which we built stochastic graph pyramids. The same observation resulted from the two different random number generators (Matsumoto and Nishimura, 1998; Mehlhorn and Näher, 1999). Bi-directed grid graphs<sup>3</sup> of size 10,000, 22,500 and 40,000 vertices, respectively, are used as the base level, which correspond to image sizes of  $100 \times 100$ ,  $150 \times 150$  and  $200 \times 200$  pixels, respectively. Statistics of all pyramids built by each specific selection (MIS, MIES, MIDES) are calculated to compare the properties of different strategies by using these parameters (see full documentation in Haxhimusa and Kropatsch, 2003):

- the height of the pyramid—*height*,
- the degree of vertices—*vertex degree*,
- the reduction factor for vertices— $\frac{|V_k|}{|V_{k+1}|}$ , and
- the number of iterations for correction (Meer, 1989)—*correction*,

The number of levels needed to reduce the graph at the base level (level 0) to an apex (top of the pyramid) consisting of one vertex is the *height* of the pyramid. The number of edges incident to a vertex, i.e. the number of non-survivors identified to the survivor represent the *vertex degree* complexity. This parameter is of importance because it is directly related to the memory costs of the representation of graphs (Jolion, 2003). The number of iterations to a complete maximal independent set for any graph in the contraction process is the iteration for *correction*. The *ratio* of the number of vertices of two consecutive levels is the *reduction factor* ( $|V_k|/|V_{k+1}|$ ). We average these parameters (sample mean  $\hat{\mu}_{\text{pyr}}$  and sample standard deviation  $\hat{\sigma}_{\text{pyr}}$ ) for every experiment (i.e. pyramid) as shown in the Table 1 over all experimental sets ( $\hat{\mu}_{\text{data}}, \hat{\sigma}_{\text{data}}$ ). The experiments are discussed in Sections 7.1, 7.2, and 7.3 and Fig. 7 summarizes the results of the first 100 of 1000 test pyramids; solid lines connect the observed number

<sup>3</sup> Note that the methods can be applied to irregular planar graphs since the second level of pyramids are irregular planar graphs.

Table 1  
Statistics on height of the pyramid, vertex degree, decimation ratios and iterations for correction

Algorithm		Height	Vertex degree			$ V_k/V_{k+1} $		Correction	
			max	$\hat{\mu}_{\text{pyr}}$	$\hat{\sigma}_{\text{pyr}}$	$\hat{\mu}_{\text{pyr}}$	$\hat{\sigma}_{\text{pyr}}$	$\hat{\mu}_{\text{pyr}}$	$\hat{\sigma}_{\text{pyr}}$
MIES	max	41.00	148.00						
	$\hat{\mu}_{\text{data}}$	20.80	70.69	4.72	3.67	2.01	1.30	3.01	0.82
	$\hat{\sigma}_{\text{data}}$	5.25	23.89	0.22	1.16	0.35	1.08	0.17	0.11
MIS	max	15.00	13.00						
	$\hat{\mu}_{\text{data}}$	14.02	11.78	4.90	0.47	2.27	0.22	4.04	1.20
	$\hat{\sigma}_{\text{data}}$	0.14	0.68	0.05	0.03	0.01	0.05	0.11	0.12
MIDES	max	13.00	18.00						
	$\hat{\mu}_{\text{data}}$	12.05	13.29	4.78	0.58	2.63	0.32	2.83	1.10
	$\hat{\sigma}_{\text{data}}$	0.40	1.07	0.13	0.04	0.10	0.16	0.15	0.10

of vertices of one particular pyramid, where the number of levels of the graph pyramid constitutes the horizontal axis and the vertical axis shows the number of vertices  $v$  in logarithmic scale at the respective pyramid level. In this choice of coordinate axes a constant reduction factor becomes a straight line. Since we reduce the pyramid to a single vertex in the apex, all the solid lines in Fig. 7 end at the  $10^0 (= 1)$ .

### 7.1. Experiments with MIS

From Fig. 7a we see that the height of the pyramid cannot be guaranteed to be logarithmic, except for some good cases. In the worst case the pyramid had 22 levels for  $100 \times 100$  vertices and 41 levels for the graph with  $200 \times 200$  vertices, respectively. Poor reduction factors are likely, as can be seen in Fig. 7, especially when the images are large. This is due to the evolution of larger and larger variations between the vertex degrees in the contracted graphs (Table 1,  $\hat{\mu}_{\text{data}}(\text{max}) = 70.69$ ). The absolute maximum in-degree was 148. The a priori probability of a vertex being the local maximum depends on its neighborhood. The larger the neighborhood, the smaller is the probability that a vertex will survive (Jolion, 2003). MIS tends to have vertices with a large neighborhood (stars). This causes the reduction factor to be very poor at highest levels, also noted in (Montanvert et al., 1991), which is mainly due to the good performance of the decimation at the first few levels where the graphs have large sizes

and a small neighborhood size (e.g. each vertex has 4 neighbors in the base level). The sample mean reduction factor of vertices is 2.01 (Table 1). The collapse of the high-order-star-like configuration into the apex causes the sudden break of the solid lines (“star-contraction” in Fig. 7a). The number of iterations necessary to complete the maximum independent set per level (iterations for correction) are approximately 3 as reported by Meer (1989). The sample mean height of the pyramid is 20.80 (Table 1). To summarize, a bounded reduction factor cannot be guaranteed and bad cases have a high probability (almost horizontal solid lines in Fig. 7a).

### 7.2. Experiments with MIES

The experiments show that the reduction factor, even in the worst case, is always better than the theoretical bound 2.0, as indicated by the dashed line in Fig. 7b. The MIES is more stable than MIS, as can be seen in Fig. 7, where the slope of the solid lines have smaller variations and never crossed the dashed line. The sample mean and sample variance of the reduction factor for MIES is smaller than in case of MIS, which implies that the height of the pyramid ( $\hat{\mu}_{\text{data}}(\text{height}) = 14.02$ ) is also smaller than that for MIS ( $\hat{\mu}_{\text{data}}(\text{height}) = 20.80$ ). The mean number of iterations for correction per level was higher for MIES (Table 1). To summarize, the reduction factor was always better than the theoretical upper bound of 2.0.

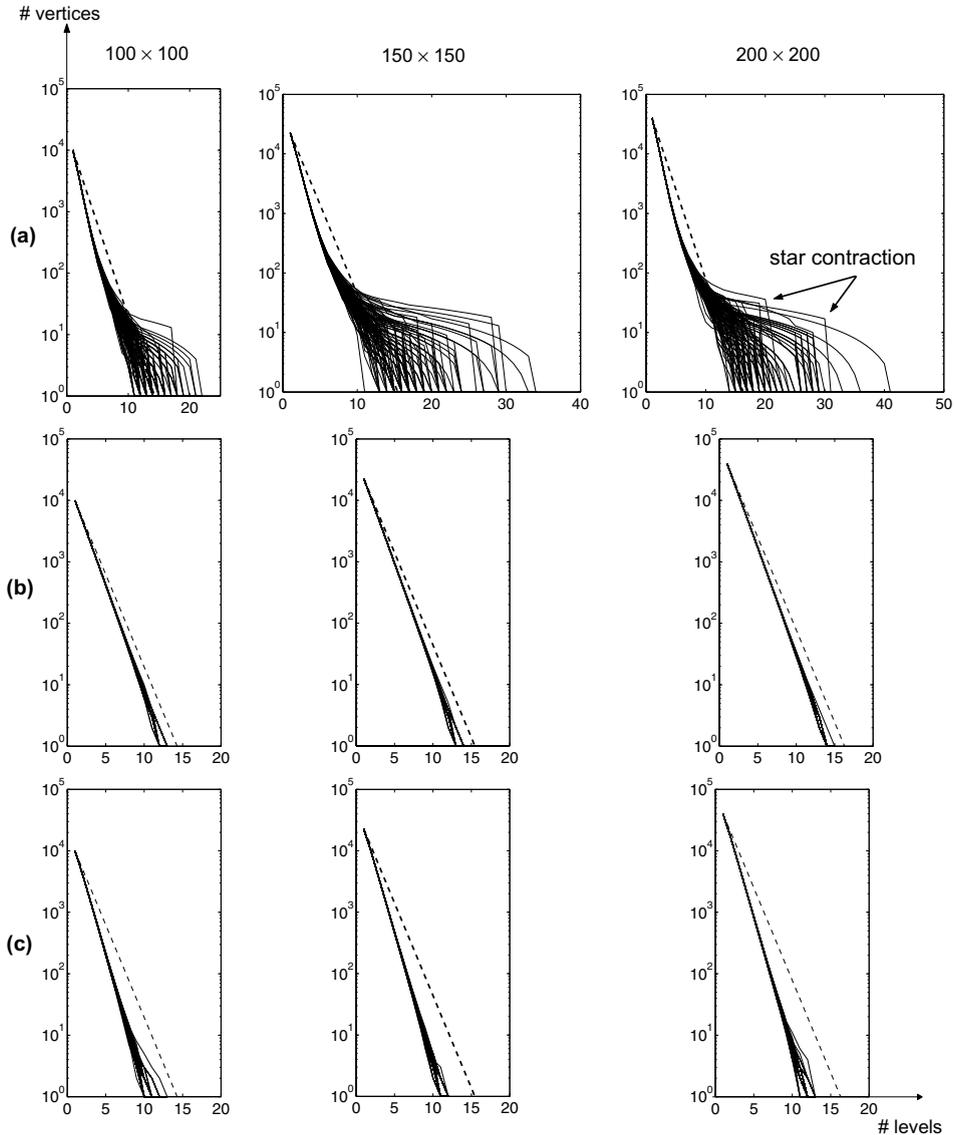


Fig. 7. Comparing reduction ratios of (a) MIS, (b) MIES and (c) MIDES. Dashed lines represent the theoretical reduction factor of 2.0.

### 7.3. Experiments with MIDES

The reduction factor is better than 2.0 (dashed line) even in the worst case as depicted in Fig. 7c. Also the maximum in-degree of the vertices is much smaller ( $\hat{\mu}_{\text{data}}(\text{max}) = 13.29$ ) than for MIS ( $\hat{\mu}_{\text{data}}(\text{max}) = 70.69$ , Table 1). For MIES and MIDES, we have not encountered nodes with a

large neighborhood as for MIS. For the case of the graph with size  $200 \times 200$  vertices, MIDES needed 13 levels in comparison to 15 levels in the worst case of MIES. The number of iterations needed to complete the maximum independent set was comparable with the one of MIS (Table 1). The MIDES algorithm shows a better reduction factor than MIES, as can be seen in Fig. 7

and Table 1 ( $\hat{\mu}_{\text{data}}(\hat{\mu}_{\text{pyr}}(\frac{|V_k|}{|V_{k+1}|})) = 2.63$ ). To summarize, the reduction factor is always better than 2.0 in all our tests. A bounded reduction factor cannot be guaranteed.

#### 7.4. Discussion of results

We have done experiments using small grid graphs  $26 \times 26$  see Jolion (2003), and encountered no significant differences in the reduction factor between MIS, MIES and MIDES. The reason for this behavior is the small size of input instances. By using graphs of larger size ( $100 \times 100$  and  $200 \times 200$ ), different properties of these selection methods occurred. The maximum degree was encountered using MIS which is why this method has problems during contractions. Thus memory costs for MIS will be higher than for MIES and MIDES. Notice however that the mean degrees are similar for all algorithms. The height stability (in the sense of smaller variation) in the first case shows that MIS, MIES and MIDES do not depend on the data. Results in Table 1 were computed using 1000 graphs of size  $200 \times 200$ . The number of iterations for correction was almost the same for all methods, MIDES gave the best reduction factors ( $\hat{\mu}_{\text{data}}(\hat{\mu}_{\text{pyr}}(\frac{|V_k|}{|V_{k+1}|})) = 2.63$ ) for all tests. MIS and MIDES have the same algorithmic complexity for the worst case. The worst case happens when the neighboring vertices have increasing random numbers. We have not encountered the worst case in our test, since it is highly unlikely. The a priori probability that a vertex survives depends on the size of its neighborhood. In Fig. 8a vertex  $u$  will be favored by MIS. Vertex  $u$  survives with the probability of  $1/2$ , since it has

only the vertex  $v$  in the neighborhood. Vertex  $v$  survives with the probability of  $1/n$ , where  $n$  is the size of its neighborhood. The center of the star will have smaller probability to survive than its leaves, which causes the poor reduction factor, since only one edge will be contracted. Since two surviving vertices cannot be neighbors, the center of the star will be pulled toward one of its leaves. But still there will be a vertex with large neighborhood. There are cases where the center of the star is the largest in its neighborhood. In these cases all its leaves will be contracted towards the star, yielding a very good reduction factor. The contraction of stars can be seen in Fig. 7, where the solid lines descent rapidly. Arrows in Fig. 7a depict examples of star contraction.

The probability that edge  $e$  in Fig. 8a will be contracted i.e. one of the end vertices will survive, is the same for all neighboring edges of  $e$  using MIES or MIDES. The existence of vertices with large neighborhoods was not encountered in MIES and MIDES, which can also be seen in Fig. 7, where there are no cases of rapidly descending lines. Since no vertex is favored, the size of receptive fields (RF) using MIES or MIDES will be better distributed, in the intermediate levels of the pyramid. There is no occurrence of very small or very big receptive fields as for MIS, where there are receptive fields as small as one pixel in higher levels of the image pyramid. Fig. 8b–d shows receptive fields of the same pyramid level of MIS, MIES and MIDES. Each vertex of the same level received arbitrary but distinguishable gray values which were propagated down to the base of the pyramid. Hence the regions with the same gray value in Fig. 8b–d correspond to the receptive

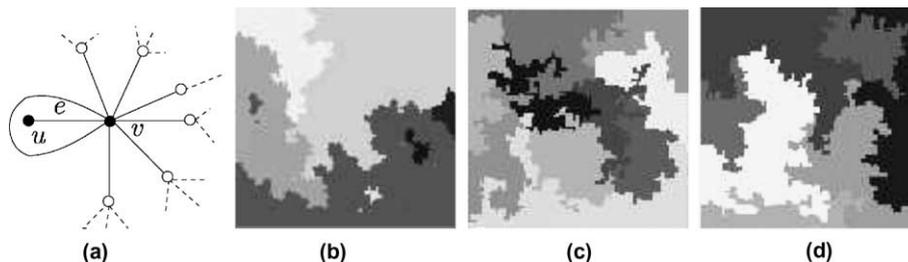


Fig. 8. (a) Vertex  $u$  is favored by MIS. (b–d) Receptive fields in the same intermediate level of the pyramid: (b) MIS, (c) MIES, (d) MIDES.

field of one vertex of the high level. Without any constraint from the data there is no need to require big variations in the sizes of receptive fields.

## 8. Eigen-image pyramids and visual attention

Pyramids can perform processes with results similar to that of the saccades of the eyes driven by visual attention. In particular the integration of a priori knowledge regarding the relevance of image regions can be realized as a top-down process with the help of a hierarchy built during a training phase. Let the saccades be viewed as a way to gain a new representation of a visual input. Instead of performing a sequential read-out of the image, an irregular pyramid structure controls resolution decrease according to the relevance of different regions. During a training phase the pyramid structure is built. It is then able to restructure new input data according to the expected relevance of different regions in the image (Langs et al., 2002; Langs and Bischof, 2002).

In Fig. 9 a scheme of the Eigen-Image-Pyramid algorithm enhancing robust PCA coefficient retrieval is given. Parallel to the training of an appearance based object recognition algorithm a pyramid structure is built. During reconstruction or recognition this structure is applied to the eigen-images as well as to the input image before robust PCA coefficient retrieval is performed to reconstruct the object. Robust retrieval determines coefficients of a particular image object by solving a set of linear equation systems based on small subsets of image pixels. The resulting hypotheses

are subject to a selection procedure that forms a recognition result with high robustness against occlusion or noise (Leonardis and Bischof, 2000).

### 8.1. Weight target contraction

The construction of the pyramid is guided by features derived from the training data (Langs et al., 2002). In the particular case while building the pyramid, survivors are selected according to the variance  $\omega_i$  of each image pixel position  $i$  over the training set (Fig. 10a). Vertices in the pyramid level  $G_0$  are assigned non-survivor or survivor status by a stochastic process (Meer, 1989). After this initialization, vertices with  $\omega_i > (1 - \tau)$ , are added to the set of survivors, where  $1 - \tau$  is the *weight target*. Consequently the first rule in Definition 2 is violated, resulting in the possibility for finer tuning during the recognition phase. Non-survivors merge with the neighboring survivor with highest correlation of the value vectors that pixels adopt in the training set (Fig. 10b), thus minimizing the loss of information. The vertex values  $\omega_{n_{i,k}}$  in a level  $G_k$  are defined

$$\omega_{n_{i,k}} = z \cdot \sum_{\omega_j: i \in \text{RF of } n_{i,k}} f(\omega_j),$$

where  $z$  is a constant ensuring convergence of contraction, and  $f(\omega_j)$  is a *weight function* controlling the distribution of receptive field (RF) sizes. Contraction is prohibited if  $\omega_{n_{i,k}}$  would exceed  $1 + \tau$ . The result is a new level  $G_1$  consisting of vertices, each representing a RF in the ground level  $G_0$ . This process is iterated and vertices in  $G_k$  successively achieve more balanced weights. Thus by

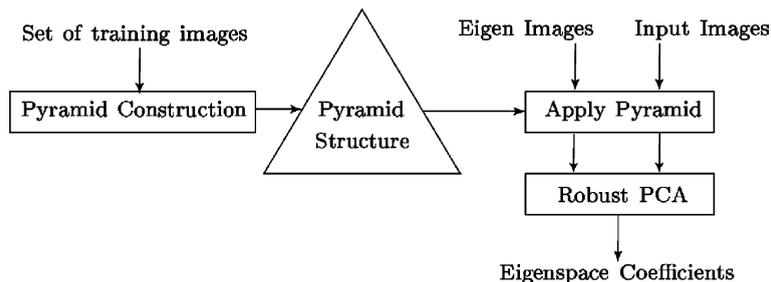


Fig. 9. The basic concept of the algorithm. It is divided into a training (left) and a reconstruction/recognition phase (right).

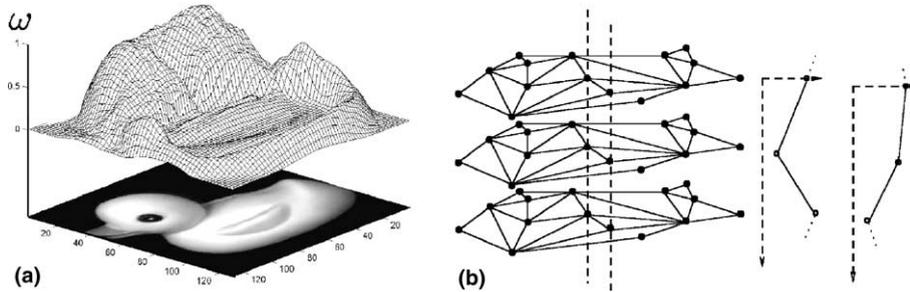


Fig. 10. (a) Weight map  $\omega$  based on variance for a training set, consisting of 36 images of a rotating duck. (b) Value profiles of neighboring vertices in the training set.

projecting new input images onto the finished structure they can be represented with different levels of detail. In contrast to regular down-sampling the structure takes the relevance and dependencies of different regions into account and increases the representation of regions with a grouping of larger weights, while preserving information if possible.

### 8.2. Enhancement of object recognition

Experiments on the images of the COIL-20 database show that applying the irregular *eigen-image pyramid* to the images improves the reconstruction error (on average by 81%) for 55% of the objects. Compared to a Gaussian pyramid the reconstruction error was improved by 61%. Experiments showed that contracting an input image by our algorithm to an optimal height with  $\sim 18\%$  number of vertices of the base level can decrease the mean squared reconstruction error down to  $\sim 53\%$  of the error achieved with full res-

olution ( $\sim 16,384$  pixels). In Fig. 11 reconstruction results of an object from images with 50% occlusion are depicted. In (a) and (b) the result after regular and irregular down-sampling are depicted, resp. In (c) the RFs of the irregular down-sampling are visualized. Note that in the apparently relevant regions the resolution is higher, resulting in a more balanced representation of the object. This supports *robust appearance based object recognition* in particular, since the hypothesis and test strategy is sensitive to input with high variation in relevance.

## 9. Conclusion

The efficiency of pyramids is tightly coupled with their ability to propagate information from any cell to any other cell in at most twice the height's steps. The gained freedom in choosing structures adapting the data may affect the height if the reduction proceeds too slowly.

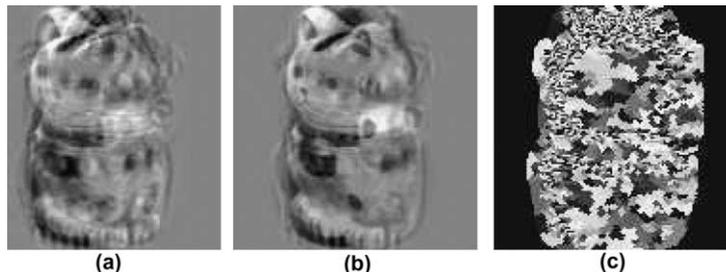


Fig. 11. Image of a cat, (a) reconstructed after regular down-sampling and (b) after irregular down-sampling (c) RFs after irregular down-sampling.

Experiments with stochastic decimation using maximal independent vertex sets (MIS) showed a problematic behavior on large images. After an initial phase of strong reduction, the reduction decreases dramatically. This is due to the evolution of larger and larger variations between the vertex degrees in the contracted graphs. To overcome this problem we proposed a method, MIES, based on matchings which guarantees a reduction factor of 2.0. As in the case of independent vertex sets, the method based on matchings does not allow the control of the directions of the contractions. The second method, MIDES, that we proposed and tested is based on directed edges and allows the control of the directions of the contractions. The experiments showed a non-decreasing reduction that was even stronger than the one obtained MIES. We have shown that 2.0 is also a bound for the reduction with MIDES if no isolated vertices are encountered. Furthermore we were able to characterize the configuration creating such isolated vertices which explains the good experimental results. The properties of these configurations are important when the random sampling is replaced by data-adaptive importance values where no prediction about statistical distribution can be made.

The final example in object recognition shows that irregular pyramids can outperform regular pyramids. Eigen-image pyramids are a way of simulating the visual attention concept by using a priori knowledge in a top-down manner. The pyramid can represent new data with different levels of detail. In contrast to regular down sampling, image regions are represented corresponding to their relevance. By choosing an optimal height of the pyramid the performance of appearance based object recognition can be improved considerably on the new representation.

### Acknowledgment

The authors would like to thank the reviewers for their constructive comments. W.G. Kropatsch and Y. Haxhimusa are supported by the FWF under grants P14445-MAT, P14662-INF and FSP-S9103-N04. Z. Pizlo is supported by

ECVision and Air Force Office of Scientific Research. G. Langs is supported by the FWF under grant P14445-MAT and P17083-N04.

### References

- Bister, M., Cornelis, J., Rosenfeld, A., 1990. A critical view of pyramid segmentation algorithms. *Pattern Recognition Lett.* 11 (9), 605–617.
- Bouman, C., Liu, B., 1991. Multiple resolution segmentation of textured images. *IEEE Trans. PAMI* 13, 99–113.
- Burge, M., Kropatsch, W.G., 1999. A minimal line property preserving representation of line images. *Comput., Devoted Issue Image Process.* 62, 355–368.
- Campbell, F., Robson, J., 1968. Application of fourier analysis to the visibility of gratings. *J. Physiol.* 197, 551–566.
- Chernyak, D., Stark, L., 2001. Top-down guided eye movements. *SMC-B* 31 (4), 514–522.
- Christofides, N., 1975. *Graph Theory—An Algorithmic Approach*. Academic Press, New York, London, San Francisco.
- Diestel, R., 1997. *Graph Theory*. Springer Verlag.
- Glantz, R., Kropatsch, W.G., 2000. Plane embedding of dually contracted graphs. In: Borgefors, G., Nyström, I., Sanniti di Baja, G. (Eds.), *Discrete Geometry for Computer Imagery, DGCI, LNCS 1953*. Springer Verlag, Uppsala, Sweden, pp. 348–357.
- Glantz, R., Englert, R., Kropatsch, W.G., 1999. Representation of image structure by a pair of dual graphs. In: Kropatsch, W.G., Jolion, J.-M. (Eds.), *2nd IAPR-TC-15 Workshop on GbR. OCG-Schriftenreihe*, pp. 155–163, band 126.
- Graham, S.M., Joshi, A., Pizlo, Z., 2000. The travelling salesman problem: A hierarchical model. *Memory Cognition* 28 (7), 1191–1204.
- Haxhimusa, Y., Kropatsch, W.G., 2003. Experimental results of MIS, MIES, MIDES and D3P. Tech. Rep. No.78, PRIP, Vienna University of Technology. Available from: <<http://www.prip.tuwien.ac.at/ftp/pub/publications/trs/tr78.pdf>>.
- Haxhimusa, Y., Glantz, R., Saib, M., Langs, G., Kropatsch, W.G., 2002. Logarithmic tapering graph pyramid. In: van Gool, L. (Ed.), *Proc. 24th German Association for Pattern Recognition Symposium, LNCS 2449*. Springer Verlag, Swiss, pp. 117–124.
- Haxhimusa, Y., Glantz, R., Kropatsch, W.G., 2003. Constructing stochastic pyramids by MIDES—maximal independent directed edge set. In: Hancock, E., Vento, M. (Eds.), *4th IAPR-TC15 Workshop on GbR in Pattern Recognition, LNCS 2726*. Springer Verlag, York, UK, pp. 35–46.
- Humphrey, G., 1948. *Directed Thinking*. Dodd and Mead, NY.
- Jolion, J.-M., 2003. Stochastic pyramid revisited. *Pattern Recognition Lett.* 24 (8), 1035–1042.
- Jolion, J.-M., Montanvert, A., 1992. The adaptive pyramid a framework for 2D image analysis. *Comput. Vision, Graphics, Image Process.: Image Understanding* 55 (3), 339–348.

- Jolion, J.-M., Rosenfeld, A., 1994. A Pyramid Framework for Early Vision. Kluwer.
- Kammerer, P., Glantz, R., 2001. Using Graphs for segmenting crosshatched brush strokes. In: Jolion, J.-M., Kropatsch, W.G., Vento, M. (Eds.), 3rd IAPR-TC-15 Workshop on GbR. CUEN, pp. 74–83.
- Koffka, K., 1935. Principles of Gestalt Psychology. Harcourt, NY.
- Kropatsch, W.G., 1995. Building irregular pyramids by dual graph contraction. IEE-Proc. Vision, Image Signal Process. 142 (6), 366–374.
- Kropatsch, W.G., Burge, M., 1998. Minimizing the topological structure of line images. In: Amin, A., Dori, D., Pudil, P., Freeman, H. (Eds.), Advances in Pattern Recognition, Joint IAPR International Workshops SSPR&SPR, LNCS 1451. Springer Verlag, Sydney, Australia, pp. 149–158.
- Kropatsch, W.G., Macho, H., 1995. Finding the structure of connected components using dual irregular pyramids. In: Cinquième Colloque DGCI. LLAIC1, Université d'Auvergne, pp. 147–158, ISBN 2-87663-040-0.
- Kropatsch, W.G., Montanvert, A., 1991. Irregular pyramids. Tech. Rep. PRIP-TR-5, Technical University of Wien, Department of Pattern Recognition and Image Processing 183/2, TU Wien, Austria.
- Kuffler, S., 1953. Discharge patterns and functional organization of mammalian retina. J. Neurophysiol. 16, 37–68.
- Langs, G., Bischof, H., 2002. Focusing visual attention in mobile robot navigation. In: Proc. 26th Workshop of the Austrian Association of Pattern Recognition, Graz, Austria, pp. 95–102.
- Langs, G., Bischof, H., Kropatsch, W.G., 2002. Hierarchical top-down enhancement of robust pca. In: Proc. 9th Workshop on SSPR&SPR, Windsor, Canada LNCS 2396. Springer Verlag, pp. 234–242.
- Leonardis, A., Bischof, H., 2000. Robust recognition using eigenimages. Comput. Vision Image Understanding 78, 99–118.
- Matsumoto, M., Nishimura, T., 1998. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans. Model. Comp. Simul. 8 (1), 3–30.
- Meer, P., 1989. Stochastic image pyramids. Comput. Vision, Graphics, Image Process. 45 (3), 269–294.
- Mehlhorn, K., Näher, S., 1999. The LEDA Platform of Combinatorial and Geometric Computing. Cambridge University Press, Cambridge, U.K.
- Montanvert, A., Meer, P., Rosenfeld, A., 1991. Hierarchical image analysis using irregular tessellations. IEEE Trans. PAMI 13 (4), 307–316.
- Pizlo, Z., 2001. Perception viewed as an inverse problem. Vision Res. 41, 3145–3161.
- Pizlo, Z., Li, Z., 2003. Pyramid algorithms as models of human cognition. Proc. IS&T Electronic Imaging, Computational Imaging, vol. 5016. SPIE, pp. 1–12.
- Pizlo, Z., Li, Z., 2004. Graph pyramids as models of human problem solving. Proc. IS&T Electronic Imaging, Computational Imaging, vol. 5299. SPIE, pp. 205–215.
- Pizlo, Z., Rosenfeld, A., Epelboim, J., 1995. An exponential pyramid model of the time-course of size processing. Vision Res. 35, 1089–1107.
- Pizlo, Z., Salach-Golyska, M., Rosenfeld, A., 1997. Curve detection in a noisy image. Vision Res. 37 (9), 1217–1241.
- Privitera, C.M., Stark, L.W., 2000. Algorithms for defining visual regions-of-interest: Comparison with eye fixations. IEEE Trans. PAMI 22 (9), 970–982.
- Rosenfeld, A., 1970. A nonlinear edge detection technique. Proc. IEEE 58, 814–816.
- Rosenfeld, A. (Ed.), 1984. Multiresolution Image Processing and Analysis. Springer Verlag.
- Rosenfeld, A., 1987. Pyramid algorithm for efficient vision. Tech. Rep. CAR-TR-299, University of Maryland, Computer Science Center.
- Rosenfeld, A., Thorston, M., 1971. Edge and curve detection for visual scene analysis. IEEE Trans. Comput. 20, 562–569.
- Stark, L.W., Privitera, C.M., 1997. Top-down and bottom-up image processing. In: Internat. Conf. on Neural Networks, vol. 4, pp. 2294–2299.
- Thulasiraman, K., Swamy, M.N.S., 1992. Graphs: Theory and Algorithms. Wiley-Interscience.
- Zeki, S., 1993. A Vision of the Brain. Blackwell, Oxford.