The Redundancy Pyramid and Its Application to Segmentation on an Image Sequence*

Jocelyn Marchadier, Walter G. Kropatsch, and Allan Hanbury

Pattern Recognition and Image Processing Group (PRIP),
Vienna University of Technology
Favoritenstraße 9/1832, A-1040 Vienna, Austria
jm@prip.tuwien.ac.at

Abstract. Irregular pyramids organize a sequence of partitions of images in such a way that each partition is deduced from the preceding one by union of some of its regions. In this paper, we show how a single pyramid can be used to encode redundant subparts of different partitions. We obtain a pyramid that accounts for the redundancy of the partitions. This structure, naturally called the redundancy pyramid, can be used for many purposes. We also demonstrate and discuss some applications for studying image sequences.

1 Introduction

Image segmentation is an important component of many machine vision applications such as object recognition and matching for stereo reconstruction. In general, segmentation techniques aim to partition an image into connected regions having homogeneous properties.

A major issue with segmentation algorithms is their stability. The partitions produced by different segmentation algorithms will be to some extent different. The same is true when a single segmentation algorithm is applied on an image sequence of a static scene under varying illumination. Comparing and merging several partitions seems an obvious way to partially solve the problem of stability.

Several techniques in computer vision and pattern recognition handle several partitions of images. A combination of different segmentations to obtain the best segmentation of an image has been suggested by Cho and Meer [2] based on the cooccurrence probabilities of points in partitions. However, they make use of small differences resulting from random processes in the construction of a Region Adjacency Graph (RAG) pyramid to generate their segmentations. Matching segmentations of different images is usually addressed as a pairwise problem, without exploiting the redundancy inherent to highly redundant images. Recently, Keselman and Dickinson [4] have proposed a method for computing common substructures of RAGs, called the lowest common abstraction. They try to find isomorphic graphs obtained from different RAGs by fusing adjacent

* This work was supported by the Austrian Science Foundation (FWF) under grants P14445-MAT, P14662-INF and S91 03-N04.

C.E. Rasmussen et al. (Eds.): DAGM 2004, LNCS 3175, pp. 432–439, 2004.

regions. While their approach is attractive, it suffers from a certain number of drawbacks. When handling real world segmentations, noise can split or merge arbitrary regions and the lowest common abstraction can not cope with these processes.

Basically these approaches try to exploit the redundancy of observations, which is widely used in robust estimation, and more generally, but implicitly, in robust computer vision techniques. In (robust) estimation, redundancy is defined as the difference between the number of parameters of a functional model, and its number of equations [3]. When the redundancy increases, the computed model is not only more precise but also more reliable [3].

Our approach is based on basic topology. We exploit the redundant structures of topological partitions. The use of this formalism guarantees that the proposed theoretical results are independent of the dimension of the space being partitioned. In section 2, we propose a set of definitions. After having recalled standard definitions in topology, we introduce new basic tools for comparing several partitions, the greatest common multiple and the lowest common divisor of partitions, whose definitions and properties are analogous to classical definitions on the set of integers. We then propose the definition of a pyramid in this framework. In section 3 we propose a fundamental theorem which enables the definition, based on these concepts, of a structure that plays a key role in the comparison of several partitions. We also propose an efficient method for constructing an approximation of the redundancy pyramid on a digital image of dimension 2. In section 4, we propose a proof of concept. The analysis of the redundancy of the structure of a segmentation of images in a sequence of moving objects in a static background leads to interesting results discussed in this section. Very redundant parts are part of a good segmentation of the background. Moderately redundant parts are moving objects, with a certain tolerance to pauses during the object's displacement. This lead to a very reliable process of background segmentation on image sequences with drastically varying illumination.

2 Basic Definitions

We recall here basic definitions from topology and propose new definitions that will help to define partitions, pyramid of partitions and the redundancy pyramid.

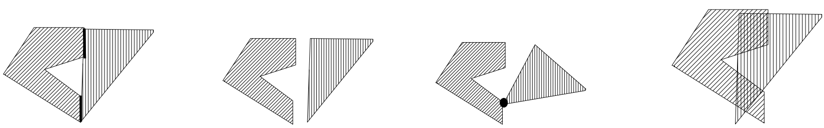
2.1 Topology

A topology on a set E is a family $\mathcal{T}$ of subsets of E (the "open" subsets of E) such that a union of elements of $\mathcal{T}$ is an element of $\mathcal{T}$, a finite intersection of elements of $\mathcal{T}$ is an element of $\mathcal{T}$, and $\emptyset$ and E are elements of $\mathcal{T}$. E equipped with a topology $\mathcal{T}$ is called a topological space.

A topological space is connected if it cannot be partitioned into two disjoint, nonempty open sets. A (topological) subspace G of a topological space E is a subset G of E such that the open sets in G are the intersection of the open sets of E with G. The complement of $F \in \mathcal{T}$ is the set $\overline{F} = E/F$. The sets $\overline{F}$ are called closed sets.

The interior $int(e)$ of a subset e of E is the largest open set contained in e. The closure $cl(e)$ of a subset e of E is the smallest closed set containing e. The boundary of a subset e of E is the intersection of its closure and the closure of its complement.

We call region a closed connected subset r of E such that $int(r) \neq \emptyset$. We define the following relations for regions. If $r \cap r' \neq \emptyset$ and $int(r \cap r') = \emptyset$ and $int(r \cup r')$ is connected, we say that regions r and r' are adjacent. r and r' are overlapping if $int(r \cap r') \neq \emptyset$. If r and r' are neither adjacent nor overlapping, we say that they are disjoint. These definitions are illustrated in Figure 1. In the example 1.c, the interesection of the two regions is composed of a single point which is on the boundary of the union. Thus the interior of their union is composed of two connected components, and we say that the regions are disjoint.



a) Adjacent regions b) Disjoint regions c) Disjoint regions d) Overlapping regions

Fig. 1. Relations between regions

2.2 Regional Covers, Divisors, Multiples, and Pyramids

We define a regional cover of a region I (e.g. the support of an image) as a set P_i of regions $r_j \in P_i$ such that two different regions from P_i are either disjoint or adjacent and $I = \cup_{P_i} r_j$. A regional cover of I is a "partition" of I into regions whose overlapping parts are thin.

We will now introduce new concepts that can be interesting when comparing several regional covers. Let P_i and P'_i be two regional covers of I. We say that P_i divides P'_i if and only if each region of P'_i has a regional cover in P_i (i.e. each region of P'_i is equal to the union of adjacent regions of P_i). We note, for convenience, $P_i|P'_i$. P_i is called a divisor of P'_i, and P'_i is a multiple of P_i. A divisor of a regional cover can be obtained by splitting its regions whereas a multiple can be obtained by merging its regions.

The least common multiple of n regional covers $P_{i,1 \leq i \leq n}$ of a region I is the multiple P of $P_{i,1 \leq i \leq n}$ such that any regional cover P'_i with $P_i|P'_i|P$ is not a multiple of one or more covers $P_{j,j \neq i}$. The greatest common divisor of n regional covers $P_{i,1 \leq i \leq n}$ of a region I is the regional cover P of $P_{i,1 \leq i \leq n}$ such that any regional cover P'_i of I with $P|P'_i|P_i$ is not a multiple of one or more regional cover $P_{j,j \neq i}$. The least common multiple (resp. the greatest common divisor) of a set of regional covers can be seen as the regional cover obtained by intersecting (resp. merging) two by two the boundaries of the initial regional covers. These definitions are illustrated in Figure 2.

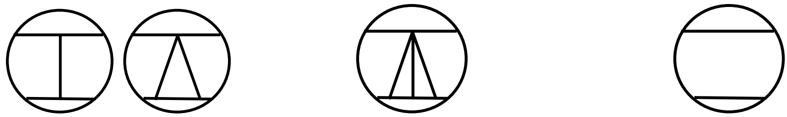


Fig. 2. The least common multiple and the greatest common divisor of two regional covers.

Irregular pyramids are well studied data structures in computer vision [5, 1]. They enable the representation of hierarchies of partitions of images. Our definition of a pyramid differs slightly from the existing ones in that we use regional covers instead of cellular partitions or graphs. This definition, although based on the same structure, leads to a simple and elegant formulation, which is expressed in a topological framework rather than in a graph framework.

We define for our purpose a pyramid $\mathcal{P}$ as a set of n regional covers $\mathcal{P} = \{L_1, ..., L_n\}$ satisfying $L_1|L_2|...|L_n$. The regional covers L_i are the levels of the pyramid, L_1 is its base level and L_n its top level. An example of a pyramid with three levels is depicted in Figure 3.

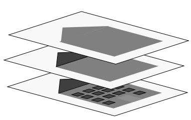


Fig. 3. A pyramid of regional covers.

3 The Redundancy Pyramid

Segmentation processes are noisy processes which can remove arbitrary regions or boundaries. The smallest common multiple of a set of covers obtained by segmentation is not stable. However, a more reliable manner to analyze common substructures of m "noisy" regional covers is to compute all the smallest common multiples of certain number i of regional covers. The smallest common multiples depend on the covers used to compute them. It then makes sense to compute their greatest common divisor L_i, which can be seen as the union of their boundary points. In this section, we will show that the L_i form a pyramid. We will give an efficient way to compute this pyramid using digital 2D images.

3.1 Definition

The following lemma simply results from the definitions. It enables one to understand how the structure of the redundancy pyramid is built.

Lemma 1. *Let $\mathcal{F}$ be a set of regional covers $P^1_{i,1\leq i\leq m_1}$. Let L_1 be their greatest common divisor. Let $P^2_{i,1\leq i\leq C^2_{m_1}}$ be all the possible smallest common multiples* *of two regional covers, and* L_2 *be the greatest common divisor of the covers* P_i^2*. Then we have* $L_1|L_2$*.*

The idea of the proof is that any intersection or difference between regions taken from different regional covers can be obtained by the union of regions from L_1. Thus regions of L_2 are equal to nonempty unions of regions of L_1 and $L_1|L_2$.

The following theorem is fundamental as it shows that the structure of the redundancy pyramid is a pyramid.

Theorem 1. *Let* $\mathcal{F}$ *be a set of regional covers* $P^1_{i,1\leq i\leq m_1}$*, and let*

- L_1 *be the greatest common divisor of* $P^1_{i,1\leq i\leq m_1}$*,*
- L_i *with* $1 < i$ *is the greatest common divisor of all the least common multiples of* i *regional covers of* $P^1_{i,1\leq i\leq m_1}$*.*

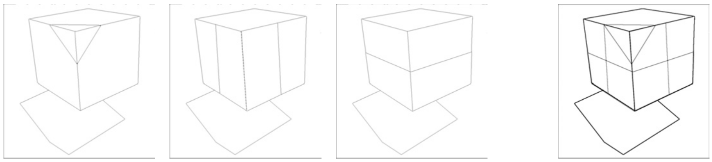
Then the set $\mathcal{P} = \{L_1, ..., L_n\}$ *is a pyramid. It is called the redundancy pyramid of* $P^1_{i,1\leq i\leq m_1}$*.*

Let us note P^i_j all the smallest common multiples of i regional covers taken from the original set, and L_i their greatest common divisor. Let P'^{i+1}_j be all possible least common multiples of two regional covers taken from P^i_j. We remark that L_{i+1} is equal to the greatest common divisor of the P'^{i+1}_j. Then we can apply the lemma 1 in order to prove the inference $L_j|L_{j+1}$. As it is true for L_1 and L_2, we have $L_1|L_2|...|L_m$. Note that by definition L_n the least common multiple of $P^1_{i,1\leq i\leq m_1}$.

3.2 Construction with Morphological Operators

The algorithm presented in this section is based on a boundary representation of each regional cover of digital images. The idea is that the set of boundary points of the level L_i of the redundancy pyramid is composed of points which are boundary points of i regional covers. Accumulating directly boundary points will not lead directly to the construction of the pyramid, as some combinations of boundary points can lead to pendant edges or isolated points. A first filtering is therefore done to remove them. On certain configurations, applying only this algorithm is not enough to filter out all undesired edges, but it produces satisfying results in most real world situations. A simple example is depicted in Figure 4. This figure shows the initial regional covers (”partitions”) of three different projected cubes, similar to the example studied by Keselman et al. [4]. The redundancy pyramid can be seen on the fourth figure, where edges have been colored according to their redundancy. The dark edges are of higher redundancy (i.e. 3), and are the common boundaries of the regions of the last level of the redundancy pyramid. The other edges have redundancies of 1, as they appear in a single image.

Although the redundancy pyramid can be built using any kind of partitions, the implementation of the preceding algorithm is straightforward when dealing with digital 2D images. The initial partitions $P_{i,1\leq i\leq n}$ are described by binary



a) Cube 1 b) Cube 2 b) Cube 3 c) Redundancy Pyramid

Fig. 4. Redundancy pyramid of images.

images (referred to as contour map in the following text) indicating the presence of contours at each point, i.e. $P_i(x, y) = 1$ if the point of integer coordinates (x, y) is a contour point for partition i and $P_i(x, y) = 0$ otherwise. Examples of contour maps are drawn in Figure 5. Typically, such images can be obtained by watershed transforms [6] or by marking contour points of labeled image partitions.

The construction of the boundary redundancy pyramid is based on an accumulation process of the contour maps. The main steps of the pyramid construction are:

- Each contour point of each contour map is accumulated in an image R of natural numbers.
- A hierarchical watershed of R is computed. We use the leveling transform of [6]. The advantage of this watershed algorithm is that ones obtain a well nested crest network and thus the pyramid in a digital form without using extra operations.

The result of this algorithm is an integer image describing the hierarchical watershed. By applying a threshold i to this image, we obtain the contours describing the ith level of the redundancy pyramid. This algorithm is not only simple but also very efficient.

4 Application to Motion Analysis and Background Segmentation on an Image Sequence

The initial data of this application is an image sequence obtained by a static camera. The captured scene can be subject to drastic illumination changes, and moving objects can occlude some parts of the static scene. A good background segmentation cannot be obtained from a single image. The main idea here is to construct initial segmentations of a certain number of images in the sequence, and to compute the redundancy pyramid of these segmentations. The low level of the pyramid will give information on the moving object, while the higher level of the pyramid will tend to segment the static scene using information merged from the sequence.

The experiment was done on a sequence where illumination varied in a way that certain images are saturated, while others are dark. On the sequence, a person is moving in front of a static background. The initial regional covers were obtained by computing the watershed of [6] on the modulus of the Deriche's gradient of the initial images, and by keeping the points not corresponding to basins. As predicted, certain regions corresponding to the sought background segmentation couldn't be retrieved correctly on all the images. They were either split or merged. Some images from the sequence and their initial segmentations are presented in Figure 5.

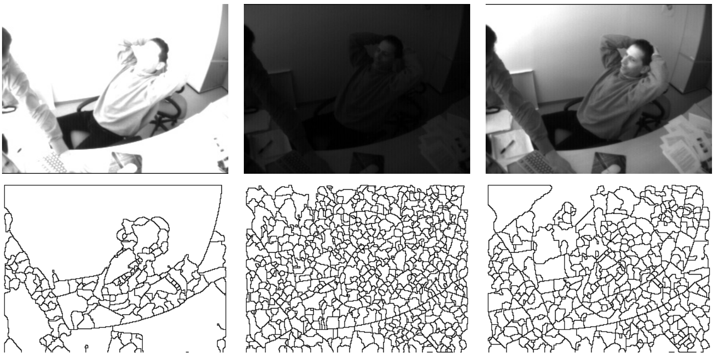


Fig. 5. Images from the sequence and their partitions (Image sequence provided by Advanced Computer Vision (ACV), Vienna).

The redundancy pyramid of the computed regional covers was computed. It is shown in Figure 6. Each image was treated in less than 2s on a laptop computer with an AMD Athlon processor at 1.8GHz. The program used was not subject to any optimization and can easily be implemented on dedicated hardware in real time.

a) Redundancy pyramid b) Level 12 c) Level 25 d) Level 40

Fig. 6. Redundancy pyramid of the image sequence

The best segmentation was obtained at an intermediate level of the pyramid. This can be explained by the fact that the contours of the background are not detected correctly on all the images. The lower levels are very noisy, which is due to the over-segmentation of the initial images. However, the trajectory of the movement can clearly be seen. The quality of the segmentations obtained at intermediate levels is outstanding, considering the initial over-segmentations used. Remark that no parameter was employed for producing the segmentations and the pyramid. The only parameter of this method is the Deriche's α which was equal to 1.5. In conditions not so extreme, the direct application of the previous method should result in stable higher levels of the pyramid. A single calibration step expressed in a number of frames would then be required in order to obtain a segmentation of a static scene of the quality as image c of Figure 6.

5 Conclusion

We have proposed new structure, the redundancy pyramid, expressed in a topological framework. We proposed an efficient algorithm in order to compute this structure on 2D digital images of partitions. It can be used in a wide number of applications ranging from segmentation fusion to generic object recognition, motion analysis and background subtraction over a sequence of images under drastically varying illumination. Some results of the last application were proposed. This validated the approach in a very complicated case. Future work include a statistical evaluation of the approach, the generalization of the algorithm to higher dimensions, to continuous images, and to images that cannot be directly superimposed on one another.

References

1. L. Brun and W. G. Kropatsch. Construction of combinatorial pyramids. In *Proceedings of the 4th IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition*, pages 1–12, 2003.
2. K. Cho and P. Meer. Image segmentation from consensus information. *Computer Vision and Image Understanding*, 68(1):72–89, 1997.
3. W. Förstner. Generic estimation procedures for orientation with minimum redundant information. 2nd Course on Digital Photogrammetry, 1999.
4. Y. Keselman and S. Dickinson. Generic model abstraction from examples. In *Proc. IEEE Conference CVPR*, volume 1, pages 856–863, December 2001.
5. W. G. Kropatsch. Abstraction Pyramids on Discrete Representations. In the proc. of DGCI2002, pages 1–21, Bordeaux, France, 2002.
6. J. Marchadier, S. Michelin, D. Arquès, Thinning Grayscale Well-Composed Images. Pattern Recognition Letters 25, pages 581-590, 2004.