

A Computer-Assisted Colorization Algorithm based on Topological Difference

Hedlena Bezerra^{1,3} Bruno Feijó¹ Luiz Velho²

¹PUC-Rio – Pontifícia Universidade Católica do Rio de Janeiro

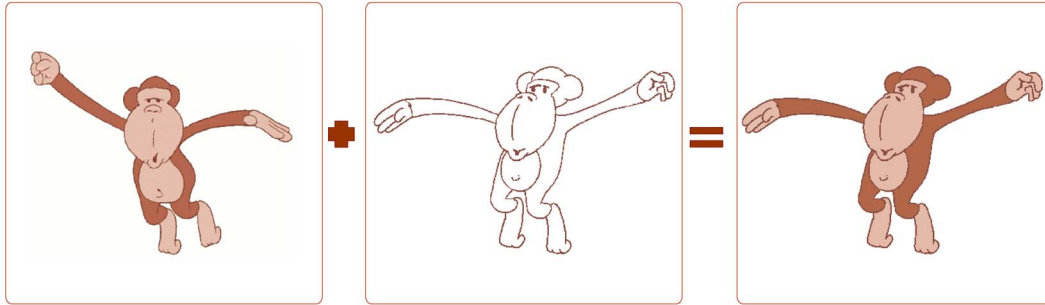
²IMPA – Instituto de Matemática Pura e Aplicada

³ARTIS GRAVIR/IMAG INRIA

Hedlena.Bezerra@imag.fr

bruno@inf.puc-rio.br

lvelho@impa.br



Abstract

Region-based approaches to cel painting typically use shape similarity and topology relations between regions of consecutive animation frames. This paper presents a new colorization algorithm based on topological differences defined over a hierarchical graph of adjacent regions, which allows an almost full automatic colorization process. Also this paper discusses other attributes that improve the solution of the image association problem.

1. Introduction

Computer Assisted Cartoon Animation is arguably one of the most challenging areas in the field of Computer Animation. Since the seminal paper of Ed Catmull [3], from SIGGRAPH 1978, where he describes the 2D animation pipeline and its main problems, there has been intense research in this area. Nonetheless, almost thirty years have passed and some of these problems remain unsolved.

There are several good reasons to explain this difficulty: On one hand, traditional cartoon animation is grounded on artistic interpretation of reality. Therefore, any computer solution has to accommodate both aesthetics results and creative control. On the other hand, 2D animation is mostly about depicting the dynamics of an imaginary 3D world through the two-dimensional strokes of the animator. In essence, these aspects together make most tasks of

computer assisted cartoon animation into very hard inverse problems, which are ill-posed and related to perceptual issues.

Ed Catmull's paper concentrated specially on the problem of inbetweening, i.e., the interpolation of key-frames in a 2D animation sequence, which is a stepstone in automating the cartoon animation pipeline.

Here we consider the dual problem of inbetweening, the tracking of 2D structural elements throughout an animation sequence. That is, we aim to establish correspondences between parts of the drawing of objects in subsequent frames. The solution of this problem has, among other things, the practical application of propagating colors from the first frame of the sequence to all the other frames. Thus, it makes the task of inking a totally automatic process. It also allows very easily to change the look of an entire animation sequence, simply by editing the rendering of one frame. Furthermore, this type of scenario integrates very well into the traditional animation process, in which the artist creates an animation drawing it frame-by-frame.

Our approach to solve the region correspondence problem is based on propagating the structural information of the drawings from one frame to the next in an animation sequence. Each frame is described by a two-dimensional graph where nodes represent regions of the drawing and arcs their adjacency relationship. Nodes have several attributes that characterize the associated regions, such as centroid, area and shape. We perform an analysis of local and global properties guided by the connectivity struc-

ture, in order to establish a partial isomorphism between the graphs in subsequent frames. Since the region correspondence problem can be inherently ambiguous, we cannot hope to solve it completely in all situations. Therefore, we employ multiple criteria and adopt strategy of only accepting a match when all criteria agree. We leave ambiguous cases for the animator to decide and create an interface that facilitates this decision. Nonetheless, our algorithm is able to establish correspondences automatically in a majority of the cases.

This paper is organized as follows. Section 2 reviews the previous works. Section 3 presents the main challenges faced in computer-assisted color propagation pointing out how to solve some of those problems using information inherent to traditional animation. Section 4 places the colorization problem in the image-based approaches. Section 5 presents how to establish the connection between a frame previously colored and another uncolored using image region information. Section 6 concludes with final remarks and a discussion of ongoing work.

2. Previous Work

Several works have been proposed to automate cel painting. Madeira *et al.* [9] use region shape similarity and Chan and Lee [4] work with topological relations. In a position-based approach to cel painting, Seah and Feng [12] work with a hierarchical motion estimation to establish similarity between frames. These previous work, however, do not consider occlusion and other topological characteristics. One of the best current works on computer-assisted colorization systems is presented by Qiu *et al* [11]. These authors propose a region-based approach using region shape similarity, topology relations, and master frames. They use a couple of predefined master frames to define pseudo 3D appearance of individual characters. The present paper does not consider pseudo 3D issues like in Qiu *et al* [11], but achieves great spatial coherence due to a robust metrics for region similarity.

3. Challenges in colorization

The first challenge in computer-assisted color propagation for traditional animation lies in the freedom practiced by the artist. In the cartoon movies of Saturday mornings, it is easy to notice that geometric or motion information are not very realistic. The animator is completely free to create any kind of situation, even the most surreal one. In addition to this, a large set of ambiguity problems might show up due to the fact that the animator's drawings and paintings are really two-dimensional projections of 3D characters as visualized in the animator's mind. In this section we will examine several scenarios and explain the difficulty of

our problem and the basic criteria involved in the association process.

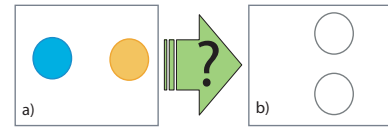


Figure 1. Ambiguity in colorization.

Though consecutive frames of an animation look very similar, sometimes animators face ambiguous situations during the colorization process. Figure 1 illustrates a situation that may occur in any animation movie. How can we determine the best association between the circumferences of images a) and b)?

Even a complete analysis of the objects movement on previous frames could not solve this ambiguity with a 100% reliability. This implies shows the difficulty to design fail-safe algorithms. However, once the frames are presented in an animation, usually, only small variations occur from one frame to the other. We show that the data concerning object localization and their topological structure in the scene can be exploited to improve on ambiguity in several situations.

Figure 2 shows a good match between objects from images in a) and b). Although this association is not guaranteed, it becomes clear that object *C* is linked to object *A*, as shown by their respective positions in the scene. Position presents a good indicator, but it is not enough to solve all cases. Figure 3 shows an ambiguous situation where the distances between probable pairs are the same.

Another clue of association between objects is shape similarity and volume. In general, as in the case of displacements, these values vary little between consecutive frames. Figures 4 illustrates an association based on geometric data. Combining position and geometric is even more effective. However, objects with identical geometry and movement, lead to the same dilemmas as in Figure 3.

To further improve topological structures of the scene should be taken into account to achieve correct associations.

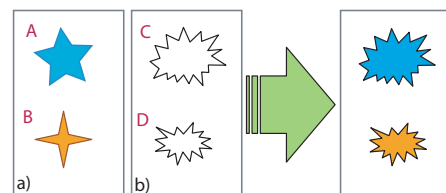


Figure 2. Association between regions by intuition.

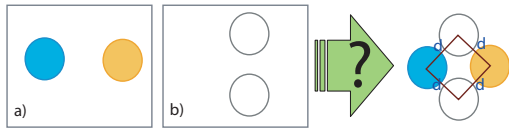


Figure 3. Distance is insufficient for coloring.

Figure 5 illustrates objects added to the scene of the Figure 3. Through the location of each object and their relative positions, it is now possible to deduce the right correspondence. In Figure 5, the relative positions of the star and square indicate the association of the circles between two consecutive frames. However, this wealth of information still does not suffice to guarantee a perfect association.

As image association is inherently an ill-posed problem, there cannot be a complete solution. Nevertheless in this paper presents an approach based on a set of local and global image attributes that allow an almost fully automatic colorization process in most of the cases.

4. Computer-assisted colorization based on images

Establishing correspondences between two different images, involves the determination of regions. The solution is given by a set of region pairs. The remaining question is, what attributes are needed to succeed in this task? This section introduces a series of local and global aspects present in animation images and shows how to apply them in our context.

The image association problem first requires a system that can identify the compound regions of the image. In outline images, each region is made up of pixels inside a boundary curve. Labeling, a common technique in computer vision, groups the pixels of an image into nonoverlapping regions based on some connectivity criteria [5]. E.g. the labeling algorithm proposed by [2], based on a flood-fill, is well adapted to find the entoured regions. In the end of the process, the image I is completely fragmented in a finite group of regions.

Once the regions of an image are identified, it is neces-

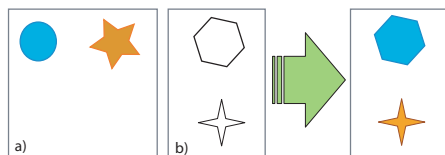


Figure 4. Association based on shape and volume.

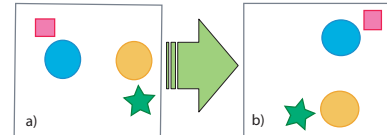


Figure 5. The topological structure of the scene can be a great ally in solving ambiguities.

sary to collect a series of attributes that will be essential to establish the associations with other regions. Some of these attributes determine local aspects, such as, area and position. These characteristics as well as topological aspects of the scene are crucial. In addition, we present other elements that can easily be extracted from any region, as well as guidelines for their application on tracking 2D animation objects. In the following subsections we will present the properties of regions, that we exploit in our matching approach.

4.1. Area

The appearance of a character can roughly be given by the shape of its open and closed lines. To transmit the idea of smooth movement, the animator makes small modifications to a previous frame, like changing the position and shape of those lines. Area can vary, but, in general, its relative change is very small [11]. Figure 6 shows two consecutive frames of an animation and that shows this claim; the general volume of the character regions is kept.

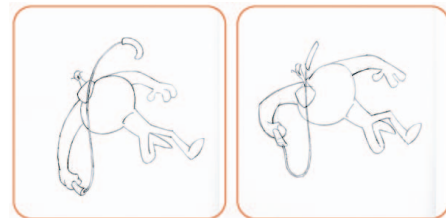


Figure 6. Small variations of local areas between two consecutive frames. Image by Marão [10]

This property makes area a useful element. After segmentation, each pixel of a same region will have exactly the same label. So, it becomes easy to compute the area (or volume) of a region by counting its pixels.

4.2. Position and contour

Maybe the most fragile attributes of a region are its position and contour. In traditional animation there is a special

connection between those two attributes. When an animator performs a drastic change in the position, the contour remains mostly the same (and vice versa) in order to make the audience follow the movement. This situation is illustrated in the Figure 7, where although the position of the whip's end changes very much, its shape remains the same. It is thus possible for the audience to recognize and track the object.

To apply this idea to regions, we associate a position to them which is given by the centroid. The region centroid can be calculated by the arithmetic average of positions (x, y) of each associated pixel. However, this might result in a point outside the region. This would not be very representative for the region's position. We realized that in practice these cases are rare for line drawings, thus we can allow a simple strategy and choose an arbitrary point.

Due to the vulnerability of this attribute, it is important to consider a threshold when calculating the proximity between two regions in different frames. This approach can also be improved by considering the movements in the preceding n frames. Thus, a translation during the last n frames can be detected and thus the position of the new centroid in frame $n + 1$ estimated (but not stated).

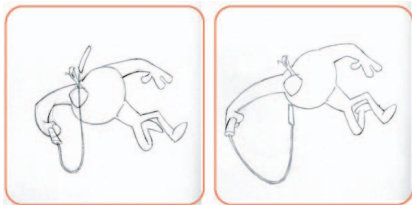


Figure 7. The variation of position is large, the variation of shape is small. Images by Marão [10].

In order to deal with contour, it is necessary to use robust algorithms to recognize shapes by some particularities that remain unaltered during a frame sequence (through which we still can recognize a form). Belongie *et al.* [1] present a simple and strong algorithm to detect associations between two distinct shapes. In their approach, shapes are represented by a group of points from internal or external objects contours. An important aspect described in the algorithm by Belongie *et al.* is the concept of shape context, which describes the allocation of points from a certain shape based on the position of a pre-determined one. Finding the association between two different shapes means determining which pair of points on both shapes share the closest shape context. Once this association is established, it is extended to the whole shape, estimating an alignment between both the first shape and the second one. The proposed method is independent of scale and object transla-

tion, and tolerates geometric distortions as well as occlusions. Shapes indicators represent a powerful assistance to the connection of different regions.

4.3. Adjacency and the proposed algorithm

Although the area, position and contour of objects are good indicators of a region, there are times when the key to a better association is the relation of each region to its adjacencies [11]. The connection between objects are also a powerful attribute to make the audience understand and follow the character's appearance and movement.

These proximity associations can be represented by adjacency graph, illustrated in Figure 8. Here, vertices represent regions, while their edges correspond to the proximity associations. When there is an edge linking two vertices on the graph, it means that corresponding regions of the image share a common boundary. Adjacency graphs represent topological region data that can be used as an indicator of good associations between images. In this way, it is important to establish the connection between the graph of a frame previously colored and another uncolored one.

There are several ways of associating two different graphs. The most simple, is also the most costly and consists in a combinatorial solution testing all pairs of nodes. This method is ineffective with images containing several regions. We present a hierarchical approach that deals with objects of the domain in an incremental way providing a more effective and satisfactory solution.

The graph from the scene must be developed in a way that permits a complete tracking of its objects. In animation frames, the background has an important role in the association process (and can be detected easily, e.g. as the region containing pixel $(0, 0)$). Even though the region seems to be distant and disconnected from the group, we will include it in the graph. This way, all regions are spanned by a single graph and it is easy to associate the root node of two consecutive frames. This allows us to gain speed, as we can now apply a hierarchical combinatorial method.

Given two associated nodes (in the first step these are the roots), we test all children against each other. Having

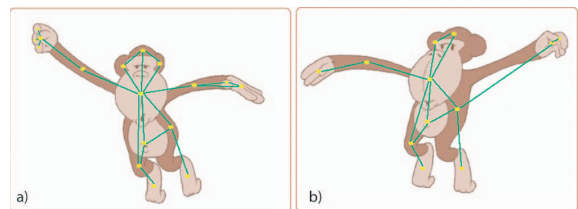


Figure 8. Adjacency graph. Regions (circles) and their adjacencies (lines) represented on both source (a) and target images (b).

found the best match we furtheron apply our algorithm recursively to the these newly associated regions. During each recursion, some nodes may not get colored in the target graphs. This occurs when the number of children nodes in the source graph is smaller than in the target graph. The algorithm finishes once every node in the source graph has been analyzed.

We have not explained yet how we test the child nodes for proximity. This is where we use our new metric that encodes topologic similarity.

We define for a region r its *adjacency function*, with respect to the morphology of its surrounding regions. We will then compare two regions by comparing the similarity of their respective adjacency function and choose the best matching pair. The adjacency function of a node is defined as follows:

Let A be a group of regions in an image (Figure 9a), let G be the graph that describes the topology of A (Figure 9b), let $V \subset A$ be a group of regions that are adjacent to a region $r \in A$. V is illustrated by dashed lines in Figure 9c). Let P_i be the set of points of the contour that defines a region i (Figure 9d).

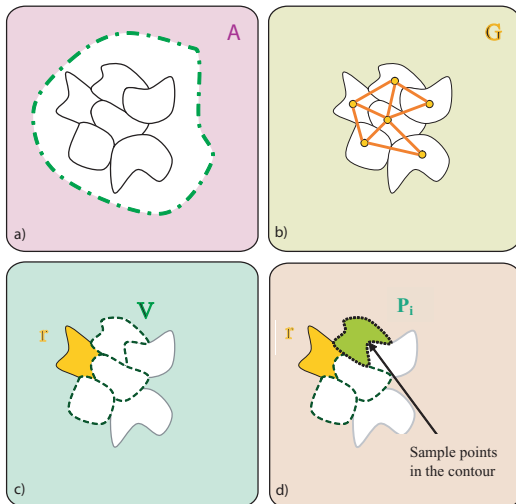


Figure 9. a) Group of regions A . b) Graphs of A . c) Adjacencies of r . d) Points located on the regions contour on r adjacencies.

Also let a be an angle and we define d as the most distant intersection of the contour P_i with a ray forming an angle a with the x-axis and origin in the centroid of the region r (Figure 9a). Then, the adjacency function $F_{(r,i)}$ of the region r with relation to its adjacent region i is the group of pairs (a_j, d_j) , where $0 < a_j < 2\pi$ (Figure 10b and Figure 11).

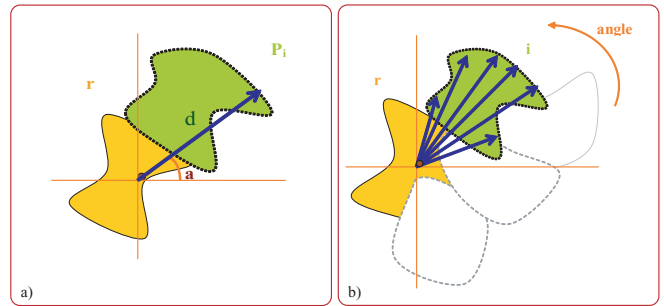


Figure 10. a) Line from the centroid of r , where a is the angle and d is the most distant intersection point with contour P_i . b) Group of lines from region r to its adjacent region i .

To measure the topological difference between regions s and t we will define the *degree of topological difference* $DTD(s, t)$. It will give us an information about the proximity of two regions based on a comparison of their adjacency functions. The distance of two adjacency functions is given by their l^2 - norm. Given two regions we want to measure the best possible association of adjacency functions. Thus the DTD is given by the sum of the distances of the best association of adjacency functions. Figure 12 illustrates the calculation of the $DTD(C, a)$, where the pair #6 and the pair #3 have the lowest value of square error. A low value of DTD indicates a good match of two regions.

The proposed algorithm calculates the degrees of topological difference for all possible relationships between regions in the current iteration level of the adjacency graph. Figure 13 shows these calculations and indicates the best pairs of association (the ones with the lowest value of DTD).

5. The colorization process

So far we have developed several criteria that we combined to a complete algorithm allowing automatic colorization of an animation sequence. This became possible due to

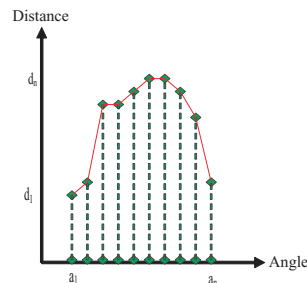


Figure 11. Adjacency Function $F_{(r,i)}$.

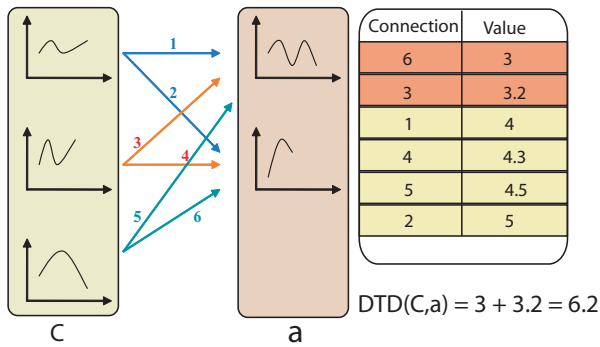


Figure 12. a) Adjacency Functions to hypothetical regions C and a . Table indicates the values of square error between each possible pair. Pairs with lower values are chosen to illustrate the difference of topology among regions C and a .

the analysis of local and global characteristics existing in a region such as: area, position, contour and adjacencies. The key to this algorithm is the intent of coloring regions with maximum precision, in other words, the association of colors between regions will only be possible if all other aspects agree.

This way, it is necessary that all uncolored regions present similar volume, contour, position and adjacencies. If any of these items are divergent, the algorithm, will not color the region. It is important to know that once the region was colored in a wrong way, this flaw can interfere on next colorizations lowering association quality. Therefore, the colorization process must not be taken with the risk of a wrong association. It is easier to the animator to color some blanked regions than undo the process. As a suggestion to get better results, the animator can be consulted, at every method discrepancy, in order to solve the dilemma. After this intervention, more regions can be colored with conviction. Still, it is possible to improve the final association quality trough the analysis of each frame examining transformations that may occur in each re-

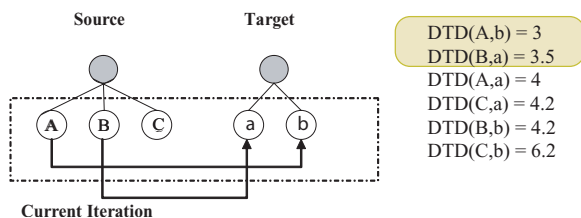


Figure 13. Statistics of association between regions from current iteration. Only the best scores will be used.

gion.

The algorithm developed in this paper is, in a certain way, a simplification of the proposed method above. To better evaluate the quality of adjacencies contribution we created a simple system only based on area, position and adjacencies. Neither user iteration nor frames examination was considered. However, even for simple approaches like this, it is possible to color almost 60% of target image, as illustrated on images 14, 15, 16.



Figure 14. a) Source image. b) Target image. Blank areas indicate new regions on the sequence.

Figure 14 shows regions from the second frame being colored using information from the first one. Even if there are some blank regions, for this example the algorithm was considered 100% successful, because these regions represent "new" regions on the image.

An example for a case where a region was not associated can be seen in Figure 15. In this case, it is possible to notice the discrepancy between the volume of this area in the original frame and in the target one. Besides that, the association was unsatisfactory as the color of the character's neck was transferred to the bottom lip. This can be explained because in the drawing, the neck disappears, whereas the bot-



Figure 15. a) Source image. b) Target image. Blank areas indicate disconnection between volumes from sample region and upcoming one.

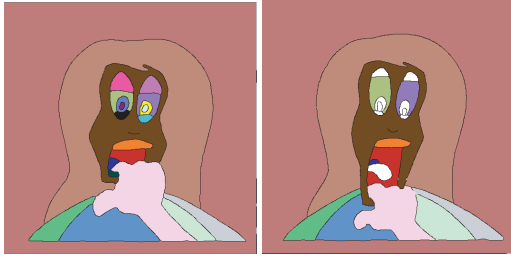


Figure 16. a) Source image. b) Target image. Blank areas indicate disconnection between volumes and adjacencies from sample region and upcoming one.

tom lip is introduced and both have quite similar appearance; the volume, positions and adjacencies seem alike.

On the third and last sequence, some blank regions occur next to the eyes and tongue. Also, on superior eyelids there is a volume variation comparing original and target images. The under eyelids and also internal regions of the eye show a big modification in their adjacencies. The uncolored tongue area is considered a new region, thus, it is not associated with any other. Nevertheless the approach correctly associated several parts, even though the images represent a drastic change.

6. Conclusions

Although the literature presents several works on computer-assisted colorization [9], [4], [11], [12], the subject has a lot of problems to be solved. The image association problem is inherently ambiguous and there is no complete solution. This paper presents a new colorization algorithm based on topological differences defined over a hierarchical graph of adjacent regions, which allows an almost full automatic colorization process. The use of the adjacency function and the Degree of Topological Difference (DTD) within a hierarchical adjacency graph generate good results. The proposed algorithm can cope with regions that come out or disappear (like a tongue that comes into view when the mouth opens or an ear that appears when the head rotates). Improvements in the quality of the proposed algorithm can be achieved by adding contour information and intelligent iteration that learns during the process. Different values can also be attributed to each element of the iteration process (area, adjacencies, position). The present paper lacks a comparative analysis with correlated works, specially with the one by Qiu [11] *et al.* Nevertheless the system we presented is quite successful despite its simplicity indicating that the new distance metric we introduced represents an important clue concerning the association of

regions. The graph structure not only accelerates computation but also helps guiding the matching process proving to be a very useful concept.

Acknowledgments. The authors thank César Coelho and Marão for all images and support as traditional animators. We also would like to thank Elmar Eisemann for unique discussions and comments regarding the present work. The authors would like to thank CNPq, CAPES, and FINEP for the financial support given to them in the form of scholarships and research contracts. Also they are thankful to the R&D centers involved in this work: VisionLab/PUC-Rio e Visgraf/IMP. The first author is supported by a grant from the European Community under the Marie-Curie project VISITOR.

References

- [1] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509–522, 2002.
- [2] H. Bezerra, B. Feijo, and L. Velho. An image-based shading pipeline for 2d animation. In *Brazilian Symposium on Computer Graphics and Image Processing - SIBGRAPI*, 2005.
- [3] E. Catmull. The problems of computer-assisted animation. In *SIGGRAPH*, pages 348–353, 1978.
- [4] C. Chang and S. Lee. Automatic cel painting in computer-assisted cartoon production using similarity recognition. *The Journal of Visualization and Computer Animation*, 8(3):285–295, 1997.
- [5] E. Davies. *Machine Vision: Theory, Algorithms and Practicalities*. Academic Press, 1990.
- [6] C. Durand. The toon project: requirement for a computerized 2d animation system. In *SIGGRAPH*, pages 285–295, 1991.
- [7] J. D. Fekete, E. Bizouarn, E. Cournarie, T. Galas, and F. Taillefer. Tictactoon: A paperless system for professional 2d animation. In *SIGGRAPH*, pages 79–90, 1995.
- [8] M. Levoy. A color animation system based on the multiplace technique. pages 65–71, 1977.
- [9] J. S. Madeira, A. Stork, and M. H. Gross. An approach to computer-supported cartooning. *The Visual Computer*, 12(1):1–17, 1996.
- [10] Marao. <http://www.maraofilmes.com.br/engole.html>, May 2006. Marão Filmes.
- [11] J. Qiu, H. S. Seah, F. Tian, Q. Chen, and K. Melikhov. Computer-assisted auto coloring by region matching. In *Pacific Conference on Computer Graphics and Applications*, pages 175–184, 2003.
- [12] H. S. Seah and T. Feng. Computer-assisted coloring by matching line drawings. *The Visual Computer*, 16(6):62–73, 2000.