Spatio-Temporal Quasi-Flat Zones for Morphological Video Segmentation

Jonathan Weber[1], Sébastien Lefèvre[2], and Pierre Gançarski[1]

[1] University of Strasbourg - LSIIT
`{j.weber,gancarski}@unistra.fr`
[2] University of South Brittany - Valoria
`sebastien.lefevre@univ-ubs.fr`

Abstract. In order to face the various needs of users, user-driven segmentation methods are expected to provide more relevant results than fully automatic approaches. Within Mathematical Morphology, several user-driven approaches have been proposed, mostly relying on the watershed transform. Nevertheless, Soille (IEEE TPAMI, 2008) has recently suggested another solution by gathering puzzle pieces computed as Quasi-Flat Zones (QFZ) of an image. In this paper, we study more deeply this user-driven segmentation scheme in the context of video data. Thus we also introduce the concept of Spatio-Temporal QFZ and propose several methods for extracting such zones from a video sequence.
Key words: Quasi-flat zones, video segmentation, segmentation personalization

1 Introduction

Following the increase of textual and then image data in personal databases and Web repositories, we are currently facing the same evolution with video data. Many video processing schemes or related use cases require a prior segmentation to get the objects-of-interest to be further processed. However, the segmentation of a given video is often not unique and depends on user needs. Thus it is necessary to rely on a segmentation method able to provide a personalized result.

Video segmentation methods designed within the framework of Mathematical Morphology may be clustered in two categories: automatic methods [2, 4] which do not require any user interaction (apart from parameter settings) and interactive methods [6–8] (but also video extension of [10]) where the user has mainly to draw some markers over objects-of-interest in order to drive the segmentation process. Results returned by automatic methods are then not adapted to user needs and often face the problem of over-segmentation. Interactive methods are more time-consuming (from a user point of view) but provide a personalized result. Another solution to solve the problem of segmentation personalization is to provide an over-segmentation which will then be reduced by the user through region merging in order him to obtain the expected segmentation. Image over-segmentation may be achieved using flat zones [11] but it then results in an extreme over-segmentation. Quasi-Flat Zones have been introduced in order to

reduce this over-segmentation while keeping interesting properties of flat zones. QFZ are based on a less restrictive criteria to build the regions, thus leading to larger regions, while keeping a low computational cost and region borders able to represent most of the frontiers between the objects-of-interest. Besides, Soille [13] notices that QFZ are not really segmentation methods but rather methods which split an image into puzzle pieces. Identifying the QFZ is then a preprocessing step in an image segmentation process based on merging of puzzle pieces. This merging may be driven by the user, thus solving the problem of segmentation personalization. Let us observe however that there is no definition of QFZ for video sequences yet.

In this article, we recall the QFZ definition in the framework of logical predicate connectivity introduced by Soille [12, 15]. We then extend this definition to video sequences, study how it can be applied to video segmentation personalization and present some preliminary results to show its interest. Finally we give some conclusions and indicate future research directions.

2 QFZ-based image segmentation

2.1 Logical Predicate Connectivity

QFZ rely on the concept of α-connected paths. A path is said α-connected if all paths between any pair of its pixels are Lipschitz-continuous, thus leading to the following definition.

Given a neighbourhood N, a path $\mathcal{P}$ made of n pixels $(p_0, p_1, ..., p_{n-1})$ is an α-connected path (α-$\mathcal{P}$) if and only if:

$$\forall i \in [0, n-2], p_i \in N(p_{i+1}) \text{ and } |f(p_i) - f(p_{i+1})| \leq \alpha \qquad (1)$$

We note α-$\mathcal{P}(p, q)$ the set of α-connected paths between p and q. This notion let us define the most simple QFZ, i.e., α-connected zones [9] which will be noted here α-CC. An α-CC is defined as:

$$\alpha\text{-}CC(p) = \{p\} \cup \{Q \mid \forall q \in Q, \alpha\text{-}\mathcal{P}(p, q) \neq \emptyset\} \qquad (2)$$

The α-CC of a pixel p is then the set of pixels to which it is linked through an α-connected path. Let us observe that flat zones are a particular case of α-CC with $\alpha = 0$. The α-CC have the following hierarchical property which will be useful later in this paper:

$$\forall \alpha' \leq \alpha, \alpha'\text{-}CC(p) \subseteq \alpha\text{-}CC(p) \qquad (3)$$

Segmenting an image into α-CC can result in an under-segmentation. If α is set too high, it will lead to a chaining effect, which may even result on a single QFZ for the whole image (it depends of course of the image under consideration and the selected α value). In order to counter this problem, new QFZ definitions based on α-CC have been elaborated. In a goal of unification of existing works, Soille and Grazzini [12, 15] have proposed a theoretical framework called logical

predicate connectivity (we recall that a logical predicate P returns true when the parameter satisfies the predicate, false otherwise). They define a new kind of QFZ (noted $(P_1, ..., P_n)$-CC here) which leads to QFZ satisfying all the n logical predicates. Various predicates may be involved, for instance: global range predicate which checks if the difference between minimal and maximal values of pixels within a QFZ is below a threshold (ω); connectivity index which is the ratio between the number of 2-pixels α-connected paths and the number of 2-pixels paths within a QFZ. This predicate is verified if the index is higher than a threshold (β). The $(P_1, ..., P_n)$-CC thus consists in seeking, for each pixel p, the largest α-CC satisfying all the predicates. Thanks to the property 3, we know that if $\alpha' < \alpha$ then α'-$CC(p)$ is a subset or equal to α-$CC(p)$. When predicates are not verified for a given value of α, we can use this property to decide to decrement α in order to check if the predicates are verified for a lower value and to loop until finding the maximal value of α for which all the predicates are verified:

$$(P_1, ..., P_n)\text{-}CC(p) =$$
$$\bigvee \left\{ \alpha'\text{-}CC(p) \; \middle| \; \begin{array}{l} \forall k \in \{1, ..., n\}, \; P_k\left(\alpha'\text{-}CC(p)\right) = true \\ \forall \alpha'' \leq \alpha', \forall q \in \alpha'\text{-}CC(p), \quad P_k\left(\alpha''\text{-}CC(q)\right) = true \end{array} \right\} \quad (4)$$

This theoretical framework is adapted to methods ensuring the unicity property. Indeed we are seeking the largest α'-CC verifying all logical predicates. It is thus not possible to consider methods which do not provide a unique QFZ segmentation. More than only a framework to unify existing definitions, the $(P_1, ..., P_n)$-CC also allows to elaborate new QFZ definitions. Three predicates are currently used within the QFZ: local range (α), global range (ω) and connectivity index (β). In the framework introduced by Soille and Grazzini, it is possible to include predicates related to other features (perimeter, area, etc.) but also to more complex descriptors (texture, gradient, etc.) as long as these predicates fulfill the condition defined in Eq. (4).

Some clues to using QFZ in multivariate images have been given by Soille [13]: α is assumed to be a vector with the same value in all components. Then, α may be easily ordered through a total ordering (decrementing $\alpha = (3, 3, 3)$ gives $\alpha = (2, 2, 2)$). Global range predicate is processed similarly, and is true only if it is verified marginally for all bands.

In the sequel of this article, we will denote by QFZ the colour QFZ built using $(P_1, ..., P_n)$-CC with only the global range predicate and a given α.

2.2 Filtering

QFZ suffer from the transition region problem. Transition regions are regions between two objects where a staircase phenomenon occurs on edge pixel values. This is due to the image discretization process and the subsequent value interpolation. This artefact leads to an over-segmentation near to the edges which will then be made of tiny QFZ. Some solutions have been proposed to solve this

problem. Soille and Grazzini [15] define transition regions as QFZ containing only transition pixels. Every pixel which is not a local extremum is considered as a transition pixel. All QFZ corresponding to transition regions are removed, and remaining QFZ are enlarged using a region growing algorithm [1]. After the removal of these regions, the amount of flat zones is reduced significantly. The solution proposed by Soille and Grazzini does not depend of any parameter and relies on a precise definition of a transition region. But from our experiments, we have observed that many regions of a few pixels remain after applying their strategy. These regions do not fit with the definition of transition regions, but are still sources of a high over-segmentation. Thus, a more efficient and robust filtering method is still lacking.

Other authors have proposed QFZ filtering methods using a QFZ minimal area thresholding step. Angulo and Serra [3] suggest to merge QFZ characterized by an area lower than a given threshold with the most similar neighbouring QFZ. With this method, no more transition region is present in the final segmentation. Zanoguera [16] removes QFZ with an area below a given threshould (thus including transition regions) before applying a Watershed transform to enlarge remaining QFZ in areas where small QFZ have been removed. Soille [14] proposes a filtering method based on an iterative increase of the minimal area, followed at each iteration by both a region growing algorithm relying on QFZ with area greater or equal to minimal area and an image simplification algorithm. The simplified image will then be segmented into QFZ at the next iteration. This process is repeated until filtered QFZ become stable. Crespo *et al.* [5] deal with flat zones and propose a flat zones merging procedure. It consists in selecting significant flat zones (according to different criteria) and expand them by incorporating most similar adjacent non significant flat zones. Following some ideas introduced in these methods, we have designed another filtering method. It relies on the Seeded Region Growing (SRG) algorithm [1] but we apply it on the QFZ rather than on the pixels (cf. fig 1). To do so, we consider a minimal area threshold similarly to existing approaches. We set all QFZ with an area greater or equal to this threshold as seeds for the SRG algorithm which is applied on the region adjacency graph. We thus obtain a much more reduced over-segmentation compared to the result obtained without filtering. The highest the area threshold is, the more reduced the over-segmentation is. But in the same time, it is much more probable to obtain an under-segmentation of some objects-of-interest. Our region growing being applied on QFZ rather than on pixels, the proposed method requires only a low computational cost.

3 Extension to video data

3.1 Limits of a 3D straight extension

The most direct extension of QFZ to video sequences is to consider a video sequence as a 3-D spatio-temporal cube. We can reuse the existing definitions, thus changing only the neighbourhood considered (spatio-temporal rather than purely spatial).
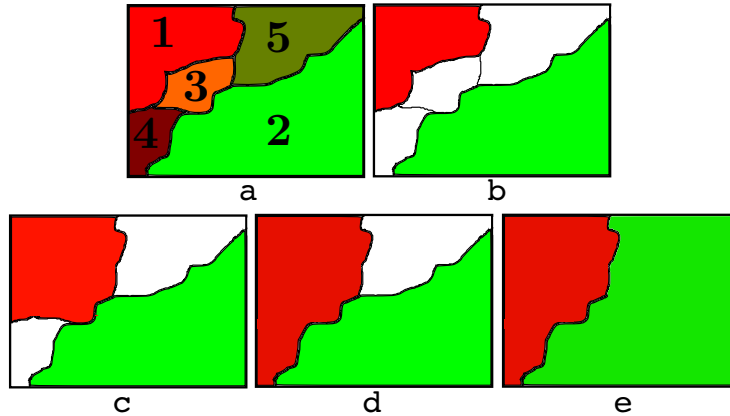
Fig. 1. Our QFZ filtering method: a) original QFZ, b) removal of QFZ with area lower than a given threshold, c) iteration #1 of SRG: QFZ 1 grows by incorporating pixels from QFZ 3, d) iteration #2 of SRG: QFZ 1 grows by incorporating pixels from QFZ 4, e) iteration #3 (last) of SRG: QFZ 2 grows by incorporating pixels from QFZ 5

Computing the $(P_1, ..., P_n)$-CC in $3D$, we obtain a higher spatial oversegmentation than in $2D$. On the 15 frames of *carphone* sample (Fig. 3.a) for $\alpha = \omega = 20$ (we will use these values in the sequel), we obtain on average $4,441 CC$ per frame in $2D$ vs. $6,779 CC$ per frame in $3D$ ($55,040 CC$ on the full sample). Indeed, by analysing the video sequences in $3D$, the considered neighborhood contains more pixels and therefore an α-CC contains more pixels (see *chaining effect* discussed above). This naturally increases the risk of violating one of the considered predicates. Thus, the largest α-CC satisfying all predicates is often produced with a small α value. This leads to tiny QFZ of only a few pixels, while such QFZ are unusable for video segmentation.

3.2 Sequential processing of spatial and temporal dimensions

As $3D$ approach is not suitable for video processing, we consider rather the $2D+t$ approach. In this approach, we successively (and no more jointly) consider the spatial and temporal dimensions, as illustrated in Fig. 2. We discuss here first the spatial to temporal $(2D + t)$ approach and then the temporal to spatial $(t + 2D)$ approach.

With the spatial to temporal approach, QFZ are first built on each frame independently. Then they are considered as nodes of a graph which are valued (here we consider the QFZ mean value). Edges are then introduced to temporally connect QFZ from successive frames and overlapping spatial coordinates. Each edge is valued by the difference between related node values. The new QFZ are the largest connected components of nodes whose connecting edges have a value less or equal to α and which do not violate any predicate. We observe that $(P_1, ..., P_n)$-CC produced significantly fewer regions in $2D+t$ ($23,926 CC$), thus
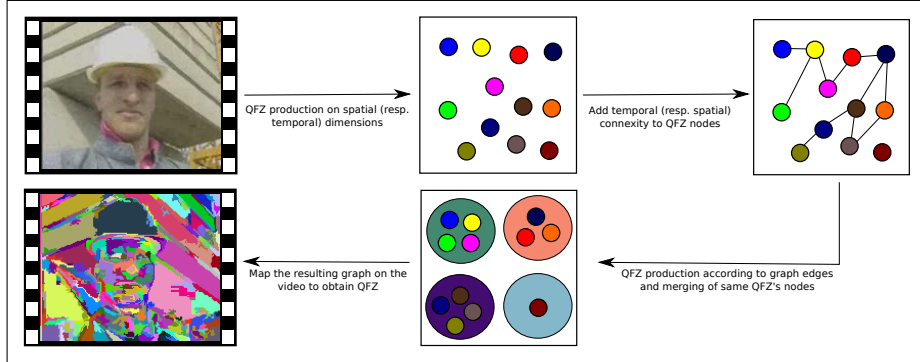
Fig. 2. Video Quasi-Flat Zones production by separated processing of spatial and temporal dimensions.

reducing the extreme segmentation we noted in $3D$ ($55,040CC$). This can be explained by the distinct computing of the two dimensions (spatial and temporal). In $2D+t$ the first processing (spatial only) produces spatially wider QFZ reducing the spatial over-segmentation. But, then the second processing (temporal) introduces temporal over-segmentation. This is due to the predicate constraints: the regions being spatially more extensive, they are less homogeneous and therefore may have significantly different mean values which will violate a predicate during the temporal computing of $(P_1, ..., P_n)$-CC.

With the temporal to spatial approach, QFZ are first built for each spatial coordinate independently, according to the temporal dimension. After this temporal processing, we therefore obtained an extreme spatial over-segmentation since, for each frame, each pixel belongs to a different QFZ. Similarly to the $2D+t$ case, we consider QFZ as nodes of a graph and apply the same process as previously, but considering here the spatial dimension instead of the temporal one. We note that $(P_1, ..., P_n)$-CC produces fewer regions ($16,830CC$) than the $2D+t$ approach due to a smaller temporal over-segmentation.

Let us observe that the $(P_1, \ldots, P_n)$-CC highlights an interesting phenomenon. Due to their different order when processing spatial and temporal dimensions, the approaches $2D+t$ and $t+2D$ induce different over-segmentations: a reduced spatial but high temporal one for the former, and a higher spatial but reduced temporal one for the latter. Nevertheless, both approaches provide better results than the $3D$ approach. Selection between $2D+t$ and $t+2D$ depends on the video under consideration. It may seem better to use the first approach with short videos of high-resolution, and to use the second for long videos of lower-resolution. Finally, let us note that the spatial and the temporal processing are both relying on $(P_1, ..., P_n)$-CC, which guarantees the uniqueness of the result. Thus they also ensure this fundamental property.

3.3 Filtering

The filtering methods presented in section 2.2 can be extended to video data. As far as our method based on a minimum area threshold is concerned, adaptation depends on the chosen approach. For the $3D$ approach, we could trivially extend the method and no longer consider a minimum area but a minimum volume. However, if considering a minimum volume would be effective in the context of truly three-dimensional images, it is not suitable for videos that are spatio-temporal and not purely spatial. Indeed, assuming a minimum volume, a QFZ having few pixels in a spatial area, but on many frames, would be kept despite the fact it is probably not an object but a part of an object. Thus, we use a threshold of minimum mean area, the mean area being computed as follows:

$$A_{mean} = \frac{\text{\# QFZ pixels}}{\text{\# frames where the QFZ is present}} \qquad (5)$$

We filter the QFZ from approaches $2D + t$ and $t + 2D$ in the same manner as for the $3D$ approach, once the QFZ building is achieved.

Similarly to the image filtering, the oversegmentation is here strongly reduced. Indeed, setting the threshold of minimum mean area to 10 pixels, we get $686CC$ for $3D$, $845CC$ for $2D+t$ and $378CC$ for $t+2D$. In addition, we get few or no under-segmentation, which makes obtained QFZ relevant for segmentation.

Filtering by minimum area threshold is very effective in reducing QFZ over-segmentation in video sequences. We obtain a very substantial over-segmentation reduction while maintaining the QFZ quality. By combining the definitions of video QFZ and the filtering area, we obtain an effective method of video pre-segmentation. This pre-segmentation can be used by QFZ merging methods to obtain user-personnalized segmentation.

4 User-driven video segmentation

Encountered in image segmentation, over-segmentation is even more present when dealing with video segmentation. For instance, when processing the sample *carphone* (Fig. 3.a) with the Predictive Watershed [4], we obtained about 2000 regions. This problem happens obviously also with segmentation by quasi-flat zones: segmenting the same sequence by $(P_1, ..., P_n)$-CC $t+2D$, with $\alpha = \omega = 20$ and a minimal area of 10 also provides an over-segmentation ($378CC$). Moreover, the resulting segmentation is not personnalized. This drawback may be solved by relying on user interaction. Such an interaction aims the user to both customize the segmentation and reduce over-segmentation. User-driven segmentation is a well-known principle in Mathematical Morphology. The main interactive morphological tool for segmentation is the marker-based watershed [10]. It has been developed for image segmentation but can be easily extended to video data. There exists also some morphological methods dedicated to video data, such as the method proposed by Gu and Lee [7] which relies on an initial interactive segmentation of a frame by marker-based watershed. The propagation of

this segmentation to following frames is performed using motion estimation and region tracking. It can be corrected by user interaction. Marcotegui *et al.* [8] propose a segmentation tool based on an initial interactive frame segmentation. This interactive segmentation is based on a multi-scale segmentation which is adapted to user needs by different tools: scale selection, region correction, region merging and marker drawing. The segmentation is propagated to the next frames via a partition projection which allows user interaction. The watershed from propagated markers method [6] is a more recent example. The first frame is segmented using marker-based watershed. Markers are extracted from the resulting segmentation and are propagated to the next frame by a motion estimation method. The next frame is segmented by marker-based watershed using the propagated markers. The user can modify the markers. In the context of QFZ, as indicated by Soille [13], a user-driven segmentation may correspond to the assembling of puzzle pieces.

We suggest a new principle for user-interactivity in video segmentation by defining a QFZ segmentation guided by markers. First, a base QFZ segmentation is produced. Then, the user draws markers on the video data. Thus, he customizes the segmentation by indicating his objects-of-interest. QFZ beneath the markers are considered as the seeds of a Seeded Region Growing algorithm [1]. The region growing will then merge the different QFZ according to their distance in terms of color, which can be related to an α parameter in the QFZ context. Since the user only see original video (and not QFZ), it is possible that several markers are found over the same QFZ. In this case, we consider that there are two possibilities: either the marker has been ill-drawn or the QFZ is ill-segmented. Here, we assume that the user has well-drawn the marker and that the QFZ has to be corrected. To do so, ill-segmented QFZ are partitioned using Seeded Region Growing with the user's markers as seeds. Thus, ill-segmented QFZ are corrected: it both improves the accuracy of the initial over-segmentation and solve the problem of having multiple markers over the same QFZ.

In order to evaluate the relevance of our proposal, we conducted some experiments on the *carphone* sequence. We compared the $(P_1, ..., P_n)$-CC $2D + t$ and $t + 2D$ methods to the Marker-Based Watershed known as the standard interactive segmentation method of Mathematical Morphology. We also compared interactive QFZ to the Seeded Region Growing involved in our method, in order to show how our method benefits from such algorithm and what it offers compared to a direct processing of Seeded Region Growing. In this perspective, we used three different settings of markers. First set of markers is just few points drawn on the median frame (Fig. 3.b). The second set is composed of heavy markers only drawn on the median frame (Fig. 3.c). The third and last set contains heavy markers on frame 4, 8 and 11 (Fig. 3.d-f). The results of these experiments are presented in Tab. 1. We denote here by precision the ratio of well-segmented pixels (affected to the good region). We used four sets of parameters (α, ω) for $(P_1, ..., P_n)$-CC to evaluate the impact of parameter settings (which have besides not been optimized). Let us observe that, whatever the markers used, our method always provides better results than the SRG.

Method	Parameters	Precision		
		(a)	(b)	(c)
$(P_1,\ldots,P_n)$-CC $2D+t$	$\alpha=\omega=10$	0.837	0.982	0.981
$(P_1,\ldots,P_n)$-CC $2D+t$	$\alpha=\omega=20$	0.841	*0.987*	0.989
$(P_1,\ldots,P_n)$-CC $2D+t$	$\alpha=\omega=30$	*0.851*	0.981	0.989
$(P_1,\ldots,P_n)$-CC $2D+t$	$\alpha=\omega=40$	*0.882*	*0.987*	0.989
$(P_1,\ldots,P_n)$-CC $t+2D$	$\alpha=\omega=10$	**0.899**	0.968	0.970
$(P_1,\ldots,P_n)$-CC $t+2D$	$\alpha=\omega=20$	0.828	0.979	0.984
$(P_1,\ldots,P_n)$-CC $t+2D$	$\alpha=\omega=30$	0.814	**0.988**	**0.993**
$(P_1,\ldots,P_n)$-CC $t+2D$	$\alpha=\omega=40$	0.837	*0.986*	0.988
Marker-Based Watershed		0.851	0.985	0.990
Seeded Region Growing		0.802	0.806	0.824

Table 1. Comparison of pixel segmentation precision with different markers (a) a few points on the median frame, b) heavy markers one the median frame, c) heavy markers on three frames)

This shows that our method by applying SRG on QFZ is better than a direct computation of SRG on pixels. Concerning the comparison to marker-based watershed, the marker-based $(P_1,...,P_n)$-CC is able achieve better results on each set of markers but only with appropriate parameter settings. Moreover, there is no single parameter combination which achieves better results with all sets of markers. But, if we exclude the first set of markers (few points in the middle frame) which are clearly insufficient, we observe that $(P_1,...,P_n)$-CC $t+2D$ and $\alpha=\omega=30$ produces the best results with both remaining sets of markers. However, like the other two interactive methods, the interactive QFZ segmentation is very sensitive to the markers given by the user. This is illustrated by Fig. 4 which shows the results obtained by marker-based $(P_1,...,P_n)$-CC $2D+t$ with parameters $\alpha=\omega=40$ and $A_{mean}=10$ with the different sets of markers.

We also compared our method to a more recent method, the Watershed from Propagated Markers [6], for which we used the binding of markers and the region-based motion propagation. As the objective here was to compare our method to a recent interactive approach in similar conditions (here the time required for the user), we did not allow the marker correction by the user and marked only the first frame. The results are given in Tab. 2 and show that marker-based $(P_1,...,P_n)$-CC in these conditions can be more accurate than the Watershed from Propagated Markers depending on the parameter settings. It implies that even if marker-based $(P_1,...,P_n)$-CC do not use motion information, its computing by assembling spatio-temporal puzzle pieces allows to obtain results comparable to other methods using motion estimation (e.g., [6]).

5   Conclusion

In this paper, we proposed both an extension of QFZ to video sequences and an interactive method for assembling these QFZ in order to build a user-personnalized
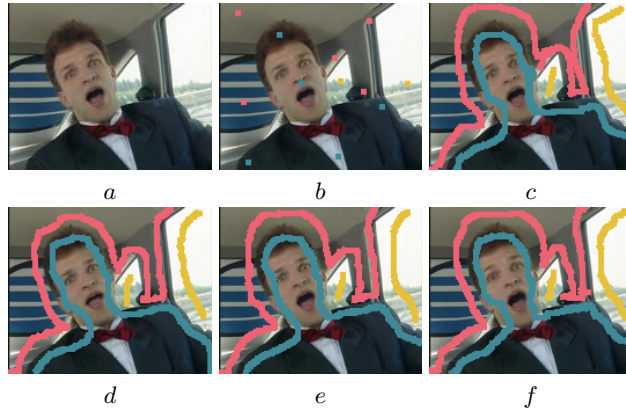
Fig. 3. Sets of markers on sequence *carphone* (a) frame 8 of extract from *carphone* sequence, b) few points on frame 8, c) heavy markers on frame 8, def) heavy markers on frame 4,8 and 11)
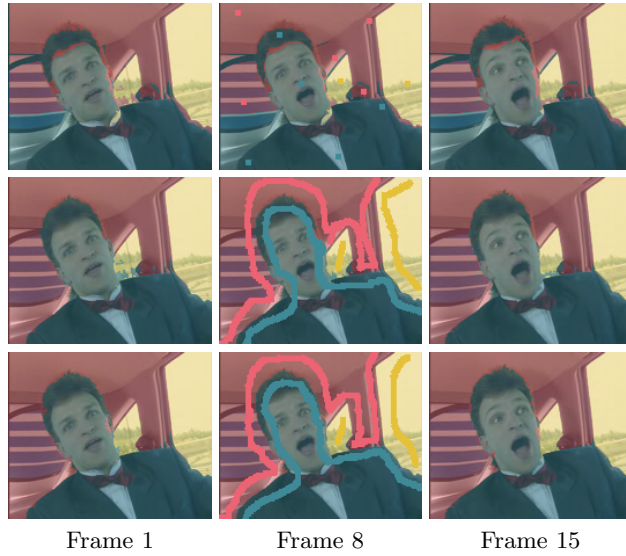

Frame 1   Frame 8   Frame 15

Fig. 4. Results of marker-based $(P_1, ..., P_n)$-CC $2D+t$ ($\alpha = \omega = 40, A_{mean} = 10$) using different set of markers, (top) points on frame 8, (middle) heavy markers on frame 8, (bottom) heavy markers on frame 4,8 and 11

segmentation. The separate processing of spatial and temporal dimensions improves the segmentation compared to a straight three-dimensional processing of video data. The proposed method for assembling QFZ according to the user's needs is intuitive and provides good results compared to other existing methods.

Method	Parameters	Precision
$(P_1, \ldots, P_n)$-CC $2D + t$	$\alpha = \omega = 10$	*0.985*
$(P_1, \ldots, P_n)$-CC $2D + t$	$\alpha = \omega = 20$	*0.984*
$(P_1, \ldots, P_n)$-CC $2D + t$	$\alpha = \omega = 30$	0.978
$(P_1, \ldots, P_n)$-CC $2D + t$	$\alpha = \omega = 40$	0.979
$(P_1, \ldots, P_n)$-CC $t + 2D$	$\alpha = \omega = 10$	0.964
$(P_1, \ldots, P_n)$-CC $t + 2D$	$\alpha = \omega = 20$	0.966
$(P_1, \ldots, P_n)$-CC $t + 2D$	$\alpha = \omega = 30$	*0.983*
$(P_1, \ldots, P_n)$-CC $t + 2D$	$\alpha = \omega = 40$	**0.988**
Watershed From Propagated Markers		0.983

Table 2. Comparison of pixel segmentation precision between $(P_1, ..., P_n)$-CC and watershed from propagated markers

Our future work will focus on a better validation of our method. We have presented preliminary results in this paper, but the method needs a deeper evaluation on a larger dataset to assess its validity. We will also focus on improving the markers. Indeed, the video is currently only marked before the processing. However, it seems relevant to be able to correct markers (like what is done in [6–8]) in order to iteratively improve the segmentation. Indeed the eventual correction of some QFZ will improve the over-segmentation at each iteration. Moreover, most of the computational cost of our approach is linked to the initial QFZ segmentation: the marker-based QFZ merging is very efficient because it is performed on the QFZ adjacency graph. We also consider to apply video QFZ on other data spaces, such as optical flow values instead of pixel values. Moreover, we plan to improve the QFZ merging process by using other features than only the color mean value. Finally, a comparative study must be achieved to determine optimal parameter settings (α, ω and A_{mean}) able to produce good results on a wide variety of video sequences.

Finally, as our method is based on a graph reduction process, we would like to design a machine learning scheme to understand how to perform segmentation from this reduction process. The idea is here to perform first a learning of some videos marked by the user and then to enable the system to segment unmarked, but simply over-segmented with QFZ, video sequences.

Acknowledgements

References

1. Adams, R., Bischof, L.: Seeded region growing. IEEE Transanction on Pattern Analysis and Machine Intelligence 16(6), 641–647 (1994)

2. Agnus, V.: Segmentation spatio-temporelle de séquences d'images par des opérateurs de morphologie mathématique. Ph.D. thesis, Université Louis Pasteur, Strasbourg (2001)
3. Angulo, J., Serra, J.: Color segmentation by ordered mergings. In: Proceedings of the IEEE International Conference on Image Processing. pp. 125–128 (2003)
4. Chien, S.Y., Huang, Y.W., Chen, L.G.: Predictive watershed: a fast watershed algorithm for video segmentation. IEEE Transactions on Circuits and Systems for Video Technology 13(5), 453–461 (May 2003)
5. Crespo, J., Schafer, R.: The flat zone approach and color images. In: Serra, J., Soille, P. (eds.) Mathematical morphology and its applications to image processing, pp. 85–92. Kluwer Academic Publishers (1994)
6. Flores, F., Lotufo, R.: Watershed from propagated markers: An interactive method to morphological object segmentation in image sequences. Image and Vision Computing 28(11), 1491–1514 (2010)
7. Gu, C., Lee, M.C.: Semiautomatic segmentation and tracking of semantic video objects. IEEE Transactions on Circuits and Systems for Video Technology 8(5), 572–584 (1998)
8. Marcotegui, B., Zanoguera, F., Correia, P., Rosa, R., Mech, R., Wollborn, M.: A video object generation tool allowing friendly user interaction. In: IEEE International Conference on Image Processing. vol. 2, pp. 391–395 (1999)
9. Nagao, M., Matsuyama, T., Ikeda, Y.: Region extraction and shape analysis in aerial photographs. Computer Graphics and Image Processing 10(3), 195–223 (1979)
10. Rivest, J.F., Beucher, S., Delhomme, J.: Marker-controlled segmentation: an application to electrical borehole imaging. Journal of Electronic Imaging 1(2), 136–142 (1992)
11. Serra, J., Salembier, P.: Connected operators and pyramids. In: Proceedings of SPIE, Non-Linear Algebra and Morphological Image Processing,. vol. 2030, pp. 65–76 (1993)
12. Soille, P.: On genuine connectivity relations based on logical predicates. In: Proceedings of the 14th International Conference on Image Analysis and Processing. pp. 487–492. IEEE Computer Society, Washington, DC, USA (2007)
13. Soille, P.: Constrained connectivity for hierarchical image partitioning and simplification. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(7), 1132–1145 (July 2008)
14. Soille, P.: Constrained connectivity for the processing of very-high-resolution satellite images. International Journal of Remote Sensing 31(22), 5879–5893 (2010)
15. Soille, P., Grazzini, J.: Constrained connectivity and transition regions. In: Proceedings of the 9th International Symposium on Mathematical Morphology and Its Application to Signal and Image Processing. pp. 59–69. Springer-Verlag, Berlin, Heidelberg (2009)
16. Zanoguera, F.: Segmentation interactive d'images fixes et de séquences vidéo basée sur des hierarchies de partitions. Ph.D. thesis, Ecole des Mines de Paris (2001)