Comics reading: An automatic script generation

Raulet Jérémy
LIASD
2, rue de la Liberté
93200, Saint-Denis,
France
jraulet@ai.univ-paris8.fr

Boyer Vincent
LIASD
2, rue de la Liberté
93200, Saint-Denis,
France
boyer@ai.univ-paris8.fr

ABSTRACT

With the advent of portable devices, reading comic ebooks is a popular activity. However, a simple scan of a comic page is not well adapted for portable device screens and a panel to panel reading without animations and adapted transitions is quite uncomfortable and not suitable. Moreover, applying manually transitions between each panel to script a complete comic book is a tricky task and seems impossible for a complete collection of comics. We present a model able to automatically script comics reading by using panel lines of force. Our results demonstrate that this model proposes a coherent solution for 87.2% of panels in an interactive time.

Keywords

Comics Script Generation, Comp.Vision & Image Processing, Mobile & WEB Graphics

1 INTRODUCTION

Nowadays, the number of comics novelty per year is in constant increase and reading them on a portable device is a common activity. These comic ebooks can be very different kinds, from a simple scan of a comic book to an electronic comic completely dedicated to the device screen and even a cartoon-like video.

Even if a comic especially created for a specific portable device seems to be the best solution, there is no appropriate solution to distribute them in an ebook format: other existing comics are scripted by a scriptwriter to produce input and output animations for each panel and exported to different portable devices. This work is performed in very different ways: by creating panel by panel transitions and animations using a dedicated tool [Rau11]; by creating a path in a comic page and displaying the entire page on the screen [Wan11]; or in the worst case, by creating a video of the comic.

We think that the first solution (i.e. creating transitions and animations panel by panel) is the best one to improve the reading experience without altering the content. However this solution is the most expensive and one can imagine how tricky the task is if the purpose is to process a comics library. Thus its automation is an interesting challenge both for researchers and commercial comics publishers.

In this paper, we consider panels reading and panels transitions. The panel extraction is realized as a preliminary step with for example Yamada et al. [Yam04], Tanaka et al. [Tan07] or Raulet et al. [Rau11] methods. For each panel, we aim at proposing an input and output animation based on its reading direction.

First, we present the terminology and the specificities of comics which are used to identify possible solutions. Then we present the related work on image retrieval and interest point detection considering the specific topics (i.e. panels transitions and reading). Then we propose our model based on image processing techniques. Results are provided comparing related work and our model. Finally, we conclude and propose future work.

2 TERMINOLOGY

In this section, we present the terminology used throughout this paper. Hereafter we precise the context and give our definitions but we do not attempt to provide an exhaustive study on comics. The interested reader should refer to [McC93] and [McC00] and as there is not a unique and unambiguous definition for all of these terms, one can find a part of this vocabulary on the website [Comi09]. From global to detail and according to our definitions, we also describe the noteworthy variations in comics to illustrate the wide range of possibilities.

Usually, a comic is described by a succession of **pages** composed by a set of image **strips**. These images, named **panels**, are colored or black and white and are often separated by **gutters**. Remark that since Rodolphe Töpffer in 1830, considered to be the modern comics creator, this page composition has been constrained to the artist by the publishing world.

Like Scott McCloud [McC93], we consider that a comic is a succession of panels. Each of them have their own size and form and are often surrounded by a black **border**. **Open panel** depicts panel without any borders. In case of overlapping between two or more panels the **overlapped panel** term is used. A panel frequently contains **speech balloons** and/or **captions** describing respectively the dialogue and the scene.

Even if this terminology covers american comics, manga, franco-belgian comics, graphic novels and all other styles, it is not enough to create a taxonomy of the domain. Many differences exist between these styles (see figure 1) depending on many factors: the technique used (brush, pencil…), the authors (two comics of the same author can be radically different)… Even for a given comic the visual representation of characters, scenes, places, that must be unique, may vary. Due to these variations, admissible for any comic readers, and the number of characters, it is not possible to build a comic database representing the collection of characters and uses it to describe the movement.

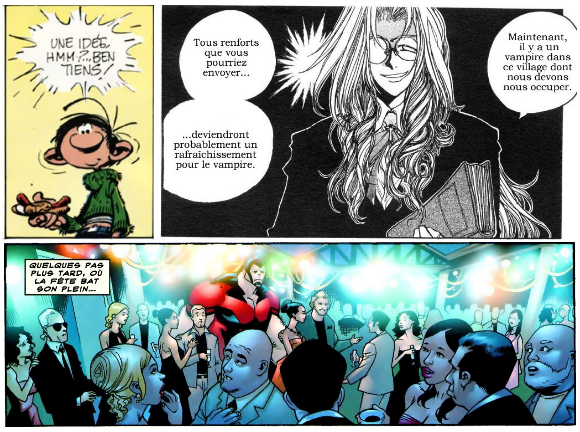

Figure 1: Top left: panel of *Gaston Lagaffe*, Top right: panel of *Hellsing*, Bottom: panel of *X-men*. These panels represent respectively franco-belgian comics, manga and american comics with different styles, levels of details and colors.

If a comics classification is not possible, one can focus on the different transitions between two successive panels and try to determine their graphics impacts.

Scott McCloud, in [McC93], has defined six forms of transition:

1. **moment-to-moment**: The second panel represents the scene a little time after the previous one, like if two photographs have been taken with a second of interval;
2. **action-to-action**: The next panel represents the next action, like a selection of key moments describing a story (see figure 2);
3. **subject-to-subject**: The same idea is illustrated in the two panels but no direct visual relation exists like in action-to-action. A common example is a phone discussion between two characters in which each panel represents a character in its own environment;
4. **scene-to-scene**: Time or distance is clearly visible between the two panels. A landscape in summer and the same in winter is an example of scene-to-scene transition;
5. **aspect-to-aspect**: The two panels describe the different aspect of the same idea or place at the same time: a beach and a character in swimsuit;
6. **non-sequitur**: No logical relation exist: suppose that figure 1 is a comic page composed by these three panels.

All of these transitions may be found in the same comic, even if the sixth is uncommon. The moment-to-moment transition is the one where panels are the most similar. But even in this case, the artist may redraw the entire panel and change, voluntary or not, a large part of it (see figure 2). It is possible that a reader does not take care about those differences but they exist. One of the most visible is the position and size of speech balloons which obscure the background.

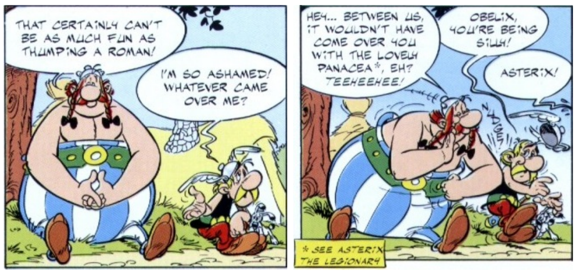

Figure 2: Page 20 of *Asterix and the Secret Weapon* panel 2 and 3 of the first strip. Excepted characters, there are many changes between these two panels. The main change is the house behind Asterix in the left panel, that disappears in the second one.

In a page, the **reading direction** is left top to right bottom, excepted in a manga. It influences the reading direction of a single panel and the eye movement should begin at the top left corner and follow a **Z pattern** in most cases.

This expected movement is disturbed by all panel elements. For example, as explained by Omori et al. [Omo04], a reader frequently skips a panel without any speech balloon. In the component hierarchical theorist, Almasy [Alm75] has explained the importance of living subjects for the reading direction. This means that the reader does not just follow a Z pattern but search important elements into the panel, like a character. Artists have several other ways to direct the reader's eyes: color contrasts, object size, level of details, closeup... In fact, comic creators determine the reading direction while generating each panel.

With the widespread of comics during the last century, artists have become accustomed to use these techniques to give the wish of pull up the reader's eyes on the second page of a double page in a comic book. They also use them to encourage the reader in turning the page after the last panel of the double page. But nowadays, panels are not necessarily arranged in a page. For example, on a mobile phone, the reader can watch each panel one by one and can have eye movement into a panel but not between two of them. Thus we need to find new techniques to direct the reader's eyes and let him concentrate on the story and not on transitions between panels.

3 RELATED WORK

To produce animation for panel transition it is necessary to detect similar contents and transformations between two consecutive panels. We have identified two main approaches:

1. Image retrieval, to detect and follow objects in a panel sequence;
2. Interest point detection and comparisons of their position to interpolate movements between panels.

Hereafter, we focus related work on these two approaches giving their advantages and drawbacks.

3.1 Content-Based Image Retrieval (CBIR)

In [Tor06], Torres et al. have explained the CBIR theory which, in particular, allows to index images with a distance function and to distinguish objects with their shape descriptors. In all CBIR methods, the main idea consists in the similarity and difference evaluation between two images.

Landré et al. [Lan07] have proposed a CBIR method using a Hamming distance and a query-by-visual-example method to compare shapes. In order to have a better perception of distances between colors, images are represented in the Lab colorspace. Then, three binary signatures per image for color, texture and shape (with a laplacian edge detector) are computed. Finally, similarities between images for each signature with a Hamming distance (a XOR binary operator) are searched. This method is well adapted to find images with the same theme (a red flower for example) and works well in general, but it is imprecise and cannot, for example, distinguish two human characters. Remark that it is possible that this method works well for a moment-to-moment transition or maybe action-to-action but it is impossible for subject-to-subject transition. Moreover, this method uses colors and some comics are "just" black and white.

The approach proposed by Fekir et al. [Fek09] is based on a Region Of Interest (ROI). This ROI is selected with a circle snake on the first image of the sequence. Then, on each image of the sequence, energies (curve consistency, gradient...) are minimized and the snake is moved. Finally, this new snake is treated like an automatic initialization on the next image and the second step (i.e. energy minimizations and snake movement) is repeated. This approach is used to follow cells in a sequence of echocardiographic images. Unfortunately, except for the moment-to-moment transition, differences between two panels are too important to implement this kind of method.

Cheung [Che07] has developed an application named MAIRE to recognize a human-like character face that helps the reader to find a particular scene in a large collection of comics. First, he has proposed the use of two CBIR methods for face detection and recommended the Adaboost one. Then, he has implemented four face recognition methods and proposed to use the EBGM (Elastic Bunch Graph Matching). These two steps enable to sort panels depending on present characters and allow the user to perform a query to find a particular scene into a large database of comics. Unfortunately, this approach requires a database of characters and as explained in section 2, it is impossible to be exhaustive. Moreover, even if it is not carefully mentioned in the paper, the detection seems to work only on full-frontal faces.

3.2 Interest Point Detection

In [Sch00], Schmid et al. have introduced two criteria for the evaluation of interest point detectors: first, the repeatability, allowing to compare the position of interest points in two images of a scene; second, the information content, allowing to measure if an interest point is really distinct one from another. They have concluded that Harris detector is the best solution for these two criteria. This method seems suitable to our problem of detecting interest points in a panel and like SIFT, SURF and ORB are posterior to [Sch00]. We present hereafter these four methods.

Gabriel et al. [Gab05] have proposed a method based on an improved implementation of Harris detector to follow an object in an image sequence. First, for each object to be tracked, a ROI is defined. Then, each ROI is described by interest points obtained from the colored version of Harris detector. Finally, the object is found in the next image with a comparison of the relative positions of interest points. The problem is that this method works with images without significant change and for any forms of transitions excepted a moment-to-moment transition we cannot initialize the ROI on each panel.

Bauer et al. [Bau07], have compared SIFT and SURF detectors. They have evaluated the invariance against rotation, scale, noise, change in lighting condition and change of view point on images of natural outdoor scenes. They have concluded that SIFT has the best performance in term of repeatability but followed very closely by SURF. They have also concluded that SURF produces fewer points and the comparison is faster. This comparison is done on photorealistic images only. We think that these methods have a bad repeatability in our case due to the precision of drawings and the difference between two similar panels. Even if our model is not based on this kind of method, we have implemented it and present benchmarks in section 5.1 to confirm our hypothesis.

Rublee et al. [Rub11], have recently presented an efficient alternative to SIFT or SURF named ORB (Oriented FAST and Rotated BRIEF). FAST is used to detect key-points and BRIEF to describe it. It seems more efficient and faster than SIFT and SURF but like [Bau07], only photorealistic images have been tested to provide benchmarks. Like SIFT and SURF, ORB is shown efficient for their experiments but has not been tested on expressive images. Our model is not based on this method but we have implemented it and present benchmarks in section 5.1.

We have presented several approaches to extrapolate a movement between two panels and no one is adequate for all transition forms. The two main problems of these approaches are:

- Methods are dedicated to follow objects in a sequence with little modification between two images;
- Methods have been evaluated only on photorealistic images.

We propose our model that enables to extrapolate a reading direction for a given comic panel.

4 MODEL

We present our model dedicated to decide both panel reading direction and panel transition. As detailed in previous work, approaches that may provide panel transitions do not exist and photorealistic approaches cannot be adapted to this kind of problem. Thus rather than a top-bottom approach providing first panel transitions to deduce the reading direction, we prefer a bottom-top approach providing first panel reading direction to deduce panels transitions. Since a comic panel is the result of an artistic process, our solution consists of determining artistic elements providing a reading direction for each panel. For that reason, our approach is based on the image processing techniques being able to collect information available in each panel. Our process is realized in 3 main steps:

1. (a) To provide a solution for any panels of any comics (i.e. colored and/or black and white), we perform an edge detection on the panel and use this information only (i.e. no color information are used hereafter);
   (b) Based on this edge detection, we extract lines of force providing a large set of possible reading information;
2. We improve our lines of force research by focusing only on dynamically defined ROI panel by panel. Thus, we keep only the most interesting part of them;
3. A classification system is finally used to determine the panel reading direction. Possible reading directions are horizontal (from left to right), vertical (top to bottom) and the two diagonals (from left to right).

Finally, according to reading directions of two consecutive panels and rules given by the scriptwriter, we provide automatically panel transitions. In practice, rules are associations between the directions and the panel transitions. These are realized independently by the scriptwriter and can be reused or changed for any comics.

4.1 Edges and Lines of Force

As a first step, we extract edges and lines of force in each panel. Lines of force is a graphical technic used since the renaissance period and are intended to convey the directional tendencies of object through space. We combine two image processing techniques to provide lines of force: an edge detection and a feature extraction technique.

As the most common edge detectors (Sobel, Prewitt, Canny) are almost interactive, we prefer the Canny detector for its detection performance [Sha02]. A Sobel kernel filter is used in the Canny detector and experiments show that a 3×3 kernel filter is the most appropriate kernel size. Other kernel sizes (i.e. 5×5 and 7×7) give a too detailed result. We follow the Canny's recommendation for the upper and lower thresholds and apply a ratio of 2:1.

Then, we use the Hough transform, as a feature extraction technique, to search the longest straight lines. We search a limited number of lines to avoid false positive with only a few lines and unfeasible results containing too many lines. This interval has been determined by a simulated annealing algorithm [Kir83] and must be in [30, 50]. These lines represent image lines of force which suggest the scene orientation. Depending on the comic style and the scene, straight lines may have different lengths.

In our algorithm (see algorithm 1), initial Canny thresholds values are used and dynamically modified according to the Hough transformation result; the Hough transform threshold is dynamically changed until the result converges to the attempted values in term of number of lines as follow: while we have not enough lines we decrease the minimal size of a straight line (Hough threshold). If the Hough threshold is too small, we decrease the Canny thresholds and repeat the Hough transform. While we have too many lines, we increase slowly the Canny thresholds. This produces a set of lines representing the panel lines of force (see figure 3).

Data: panel
Func: image `Canny` (*imageSrc, lowerThreshold, upperThreshold, SobelFilterSize*) ;
Func: setOfLines `Hough` (*imageSrc, threshold*) ;
Result: LINES (lines of force set)

thresholdCanny $\leftarrow$ 401;
minNbLine $\leftarrow$ 30;
maxNbLine $\leftarrow$ 50;
maxThresholdHough $\leftarrow \frac{3}{4}$ panelDiagonal;
minThresholdHough $\leftarrow \frac{MIN(panelWidth,panelHeight)}{10}$;
repeat
  dst $\leftarrow$ `Canny` (*panel, thresholdCanny, thresholdCanny* $\times$ 2, *3*) ;
  thresholdHough $\leftarrow$ maxThresholdHough;
  repeat
    LINES $\leftarrow$ `Hough` (*dst, thresholdHough*) ;
    thresholdHough $\leftarrow$ thresholdHough-1;
    if *thresholdHough* $\leq$ *minThresholdHough* **then**
      thresholdCanny $\leftarrow$ thresholdCanny-100;
      break;
    end
  until *nbLine* $<$ *minNbLine*;
  if *thresholdCanny* $\leq$ *0* **then**
    break;
  end
  thresholdCanny $\leftarrow$ thresholdCanny+5;
until *nbLine* $>$ *maxNbLine* ***OR*** *nbLine* $<$ *minNbLine*;

Algorithm 1: Lines of force detection.

However, the detected lines of force are, in most cases, disturbed by the border and speech balloons, so we propose a method to improve this result.

4.2 ROI

In figure 3, one can note that panel borders and speech balloons also produce lines of force. Since both are generally composed by straight lines, their impact on the line of force detection is very important. To avoid the noise generated by borders, the region on which our algorithm is applied is reduced by 10% on left, right and bottom of the panel. This value guarantees that borders will be removed and does not affect line of force detection as shown in our experimentation.

Figure 3: a) A panel of *Le donjon de Naheulbeuk*. b) Canny edges detector on (a). c) Lines of force with Hough transformation on (b). d) Latter lines on (a).

Comic artists follow some rules when creating speech balloons: they are often close to speaking character faces which are frequently located in the center of the panel; they are located where they do not hide a significant part of the drawing: for example it is uncommon that a speech balloon mask a part of a character's face; comic artists use the rule of third to place important elements in the panel and speech balloons are commonly located at the periphery. According to these principles, speech balloons are frequently placed on the top of the panel, in the sky or the scenery. In the case where speech balloons are in a particular layer, we do not consider them for the line of force detection. In any other cases we remove the upper part of the image according to the rule of third.

Figure 4 presents the line of force detection on a reduced ROI. As one can see, borders and speech balloons are not considered any more and the result is more relevant.

Now, these lines can be classified to extrapolate the reading direction.

4.3 Classification and Interpretation

We propose a classification and an interpretation system for the lines of force previously detected. We consider

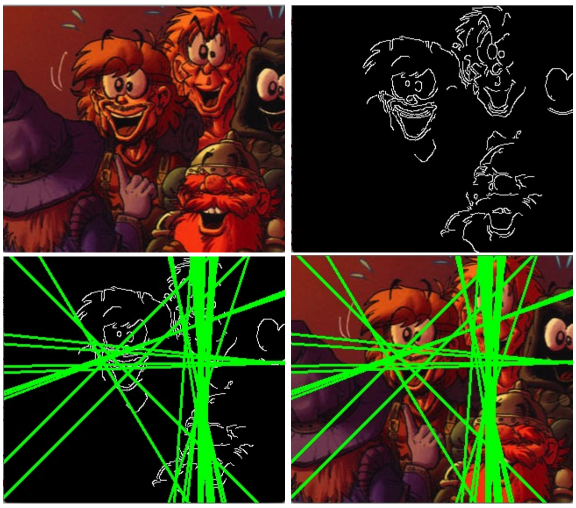

Figure 4: Figure 3 with a reduced ROI.

that the reading direction depends on the most representative direction of the lines of force.

First, we calculate the non-oriented gradient for each line. Then, lines of force are classified into 5 groups depending on their gradients: horizontal, vertical, the two diagonals and others. We fix a precision of $\pm\ \pi/16$ radian compared to horizontal and vertical axis and an angle of $\pm\ \pi/4$ radian for the diagonals. Others lines which are not classified in one of these four groups constitute the group named "other". Figure 5 presents these groups. Remark that the surfaces of the four groups are equivalent and sum to the surface of the group named other (i.e. half of the circle surface). The figure 6 illustrates this classification on the example used throughout the article.

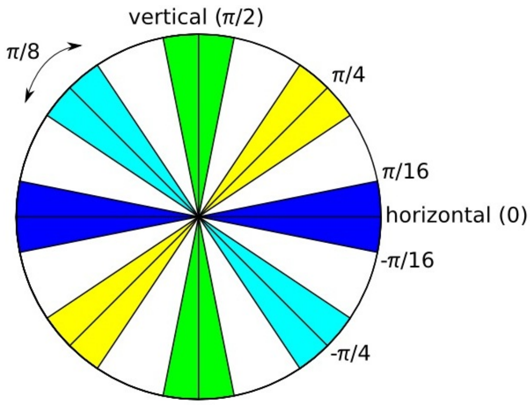

Figure 5: Lines of force classification. Each group is represented by a color and has an angular distance equals to $\pi/8$. The group named other is in white.

To select the panel reading direction, we sort all of these groups (except group other) according to the number of lines they contain. Finally, we compare the larger group to the total number of lines in two steps: first, if the group contains more than 33% of the lines, this group

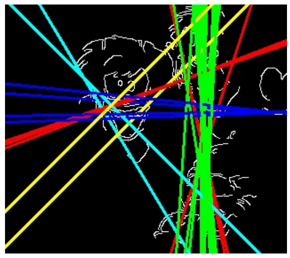

Figure 6: Each line color represents a group, the group other is in red.

is chosen to become the reading direction; otherwise, if this group contains more than 25% it is retained. Note that if the second larger group contains also more lines than the percentage that permits the choice of reading direction, it is chosen as a direction applicable if the panel is too large to be displayed in full screen (see algorithm 2 for more details).

```
Data: setoflines
Result: readingDirection optionalDirection
N ← |setoflines|;
foreach Line in setoflines do
    find gradient of line;
    classify Line according to its gradient;
    /* 5 groups: horizontal,
    vertical, two diagonals, other */
end
threshold ← N/3;
/* In all following conditions, we
do not test the group other */
sorting groups in decreasing order;
if |first group| > threshold then
    readingDirection ← group orientation;
else
    threshold ← N/4;
    if |first group| > threshold then
        readingDirection ← group orientation;
    else
        return unclassified;
    end
end
if |second group| > threshold then
    optionalDirection ← group orientation;
    return readingDirection and optionalDirection;
else
    return readingDirection;
end
```

Algorithm 2: Lines of force interpretation.

Even if our system provides a classification for the most of the panels, some of them remain unclassified (see section 5.2). These cases occur when lines of force are mainly classified in the group other or when groups are balanced. Since we decide to provide one of the attempt reading directions (horizontal, vertical and the two diagonals), for these cases, the reading direction is finally given by a scriptwriter.

Note that rules can be added or changed in our system implementation to interpret the reading direction. As an example, one might want to add a flicker animation if the two diagonal groups are the majority.

5 RESULTS

In this section, as mentioned in section 3.2, we focus first on results using SIFT, SURF and ORB. Then we present results provided by our model.

5.1 SIFT, SURF and ORB

As described above, these methods are well adapted for photorealistic image retrieval and have their own advantages and drawbacks. One can easily imagine to adapt these techniques to the comic panel transitions by finding corresponding points in successive panels. As we have evaluated carefully these methods on comic panels and have remarked that they are not dedicated to this kind of images, we provide hereafter salient results on panel transitions. However, for clarity purposes, we only focus on a single couple of representative panels where the transition is the most favorable (i.e. moment-to-moment transition).

Note that, as an implementation detail, we use OpenCV library for SIFT, SURF and ORB and we test the feature matching with flann, fern and brute force.

Figure 7 presents results of interest points matching on a panel. As one can remark, speech balloons and onomatopoeia produce noise during the interest points detection. Figure 8 illustrates the same algorithms using the ROI defined in section 4.2.

As concluded by Bauer et al. [Bau07], SIFT produce too many points to be easily readable. On top of figure 8 there are numerous good detections and matchings but also numerous false positives.

By opposition, SURF gives a very poor repeatability of detection and it is impossible to provide a confident panel transition (see middle of figure 8).

As mentioned by Rublee et al. [Rub11], ORB is a good compromise between SIFT and SURF. ORB (see bottom of figure 8) has a better repeatability than SURF but also fewer points compared to SIFT. However, here also, matching is still shoddy like for SURF and SIFT.

As mentioned in section 2 difference between two successive panels are often more important than between two photographs of the same scene (like in [Bau07] and [Rub11]). Also, since a large part to the repeatability is distorted by many changes between two panels (i.e. shapes, colors...), this kind of methods cannot be applied to comic images.

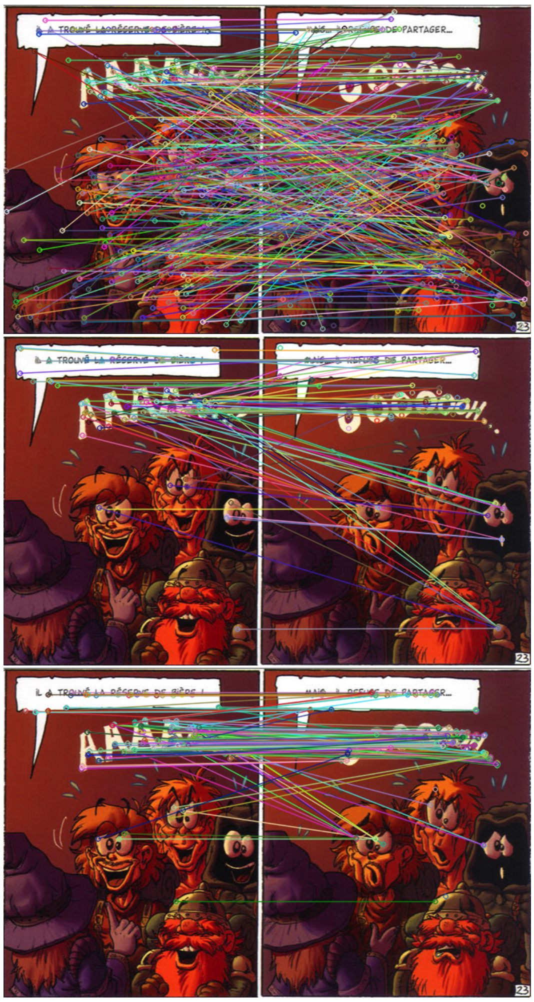

Figure 7: Up to down: SIFT, SURF and ORB detectors on two panels of comic *Le Donjon de Naheulbeuk*.

5.2 Results of Our Model

We experiment our model on a large collection of different type of comics: comic books (*X-men* and *Star Wars*), Franco-Belgian comics (*Gaston Lagaffe*, *Asterix*, *Tintin* and *Le Donjon de Naheulbeuk*) and mangas (*Naruto*, *One Piece* and *Hellsing*). They constitute a set of 2,000 panels. Our system has classified the panel set in 16 minutes with a Intel Core 2 Duo processor (2.26GHz) and 2Go RAM.

87.2% of the panels have been classified into one of horizontal, vertical or the two diagonals groups. 12.8% of panels are classified in the group other. This classification is presented in table 1; the horizontal reading direction is a majority with a distribution of 54.5%. This can be explained by the horizontal reading direction and the page format where panels are more frequently horizontal than vertical. The vertical reading direction is not well represented, which can be explained by the fact that few panels are vertically extended. Remark that our system has proposed the same reading direction (vertical) for the two panels used on figure 9.

reading direction \ optional direction		Horizontal	Vertical	Diagonal $\pi/4$	Diagonal $-\pi/4$
Horizontal	54.5%	N/A	0.9%	0.9%	1.4%
Vertical	8.5%	0.6%	N/A	0.8%	0.7%
Diagonal $\pi/4$	14.2%	1%	0.7%	N/A	0.6%
Diagonal $-\pi/4$	10%	0.5%	0.3	0.6%	N/A
Other	12.8%	N/A	N/A	N/A	N/A

Table 1: Summary of the interpretation of reading direction on 2,000 panels. Column 2 is a total for a given direction. Columns 3 to 6 correspond to the optional direction, applicable if the panel is too large to be display in full screen.

Figure 8: Up to down: SIFT, SURF and ORB detectors on ROI of two panels of comic *Le Donjon de Naheulbeuk*.

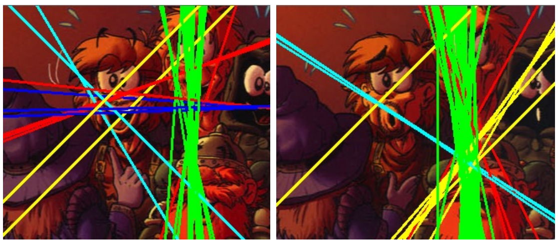

Figure 9: Lines of force of two successive panels of comic *Le Donjon de Naheulbeuk*. On these two panels, our system has proposed a vertical reading direction.

Note that only 10% of the panels (excepted panels in the group other) have an optional direction. That means that the difference between the first and the second orientation of a panel is generally large enough.

As our test protocol consists of a comparison between results provided by our model and by a panel of humans testers, we have selected randomly 100 panels in our set. Note that testers are comic readers (27%), game designers (13%), graphic designers (20%) and other (40%). We asked them to select only one group for each comic panel and we retain only the majority group. As one can see in table 2, reading directions proposed by testers and our system are very similar. The first column represents results with our tool for these 100 panels. Our results and human choice are concordant in 78% of the cases (91% for horizontal, 100% for vertical, 56% for $\pi/4$ diagonal, 46% for $-\pi/4$ diagonal and 41% for others). Remark that for horizontal and vertical, the results are identical in more than 93% of the cases. Note that, even if it does not appear in this table, the answers provided by testers for each comic panel, are generally distributed uniformly with a dominant group.

6 CONCLUSION

We have proposed a model allowing to script a comic ebook reading by adding panel reading direction and inputs/outputs animations. Our method has classified 87.2% of panels in the same way a human would do in 78% of cases. The purpose of this work does not consist of replacing scriptwriter but to suggest animations to them and to reduce the time of a comic script generation. Our comics reading system is integrated to a very complete sketch-based interface to script comics reading. In future work, we plan to use curved lines as lines of force and we aim at integrated perspective to propose more complex animations.

7 ACKNOWLEDGEMENT

This work was supported by TEKNEO (producer of entertaining applications (video games), as well as serious applications (Serious games)). Jérémy Raulet was a recipient of a CIFRE fellowship.

8 REFERENCES

[Alm75] P. Almasy. *La Photographie, moyen d'information*. Tema-éditions, 1975.

our tool \ human choice		Horizontal	Vertical	Diagonal $\pi/4$	Diagonal $-\pi/4$	other
Horizontal	48	**44**	3	1	0	0
Vertical	12	0	**12**	0	0	0
Diagonal $\pi/4$	16	2.33	4.33	**9**	0	0.33
Diagonal $-\pi/4$	12	0	3	0	**8**	1
Other	12	2	1.5	0.5	3	**5**
total of human choice		48.33	23.83	10.5	11	6.33

Table 2: Summary of the interpretation of reading direction on 100 panels. Column 2 is a total for a given direction for our tool. Columns 3 to 7 show testers majority direction choices compared to results with our tool.

[Bau07] J. Bauer, N. Sünderhauf, and P. Protzel. "Comparing several implementations of two recently published feature detectors". *Proc. of the International Conference on Intelligent and Autonomous Systems*, Vol. 6, No. pt 1, pp. 143–148, 2007.

[Che07] S. C. S. Cheung. "Face Detection and Face Recognition of Human-like Characters in Comics". Tech. Rep., City University of Hong Kong, Department of Computer Science, 2007.

[Comi09] "Comicbook Dictionary". Oct 2009. http://comicbooks.wikidot.com/comicbook-dictionary.

[Fek09] A. Fekir, N. Benamrane, and A. Taleb-Ahmed. "Détection et suivi d'objets dans une séquence d'images par contours actifs". *Proc. CIIA*, Vol. 547, 2009.

[Gab05] P. Gabriel, J.-B. Hayet, J. Piater, and J. Verly. "Object tracking using color interest points". In: *Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on*, pp. 159–164, IEEE, 2005.

[Kir83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. *Optimization by simulated annealing*. Vol. 220, Washington, 1983.

[Lan07] J. Landré and F. Truchetet. *Image retrieval with binary hamming distance*. Citeseer, 2007.

[McC00] S. McCloud. *Reinventing Comics: How Imagination and Technology Are Revolutionizing an Art Form*. HarperCollins, 2000.

[McC93] S. McCloud. *Understanding Comics: The Invisible Art*. Tundra Publishing, 1993.

[Omo04] T. Omori, T. Igaki, T. Ishii, K. Kurata, and N. Masuda. "Eye catchers in comics: Controlling eye movements in reading pictorial and textual media.". *28th International Congress of Psychology*, pp. 211–219, 2004.

[Rau11] J. Raulet and V. Boyer. "A Sketch-based Interface to Script Comics Reading". *SIGGRAPH Asia 2011 Sketches*, p. 3, Dec 2011.

[Rub11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. "ORB: an efficient alternative to SIFT or SURF". *International Conference on Computer Vision*, pp. 2564–2571, Nov 2011.

[Sch00] C. Schmid, R. Mohr, and C. Bauckhage. *Evaluation of Interest Point Detectors*. Vol. 37, Springer, 2000.

[Sha02] M. Sharifi, M. Fathy, and M. T. Mahmoudi. "A Classified and Comparative Study of Edge Detection Algorithms". *Information Technology: Coding and Computing*, pp. 117–120, 2002.

[Tan07] T. Tanaka, K. Shoji, F. Toyama, and J. Miyamichi. "Layout Analysis of Tree-Structured Scene Frames in Comic Images". In: *Proceedings of the 20th international joint conference on Artifical intelligence*, pp. 2885–2890, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2007.

[Tor06] R. d. S. Torres and A. X. Falcão. "Journal of Theoretical and Applied Informatics (RITA)". *RITA*, Vol. XIII, No. 2, pp. 165–189, 2006.

[Wan11] A. D. Wandani, G. Wee, and W. S. Moses. "Designing Interactive Mobile Comics for Multi-Touch Screen Phones". *International Conference on Future Information Technology IPCSIT*, Vol. 13, pp. 332–336, 2011.

[Yam04] M. Yamada, R. Budiarto, M. Endo, and S. Miyazaki. "Comic Image Decomposition for Reading Comics on Cellular Phones". *IEICE Transactions*, Vol. 87-D, No. 6, pp. 1370–1376, 2004.