Formulation and Use of Time-to-Collision in Content-Based Retrieval

ANKUSH MITTAL[1]
LOONG-FAH CHEONG[2]

[1]Electronics and Computer Engg.
Indian Institute of Technology, Roorkee, India
Email: ankumfec@iitr.ernet.in

[2]Electrical and Computer Engg.
National Univ. of Singapore, Singapore
Email:eleclf@nus.edu.sg

Abstract. Time-To-Collision (TTC) is a common vision feature used to avoid obstacles, and is frequently used in cinematography to induce emotional effects. In this paper, we show how TTC can lead to several high-level video categories. The extracted TTC shots (low-level feature) are mapped to their corresponding high-level indices. The information conveyed by neighboring frames or shots (i.e., contextual information) is used to facilitate the mapping process. Our proposed TTC detection algorithm is based on computing TTC from the divergence of the image velocity field. A simple and novel method known as the *pilot cue* is used to further refine our algorithm. The experiments results are promising and show the effectiveness of the approach.

Keywords: Multimedia indexing, Contextual information, Time-to-collision, Semantic indices, Content based retrieval, Pilot cue

(Received June 19, 2006 / Accepted October 05, 2006)

1 Introduction

Time-To-Collision (TTC) is the time needed for the observer to reach the object, if the instantaneous relative velocity along the optical axis is kept unchanged [1]. According to Marr [2], there exists a specific mechanism in the human visual system, designed to cause one to blink or to avoid a looming object approaching too quickly. A video shot with a small TTC evokes fear because it indicates a scene of impending collision. Thus, TTC can serve as a potent cue for the characterization of an accident or violence. Fig.1 shows the effectiveness of TTC in the content characterization with the example of a climax. Before the climax of a movie, there are generally some chase scenes leading to the meeting of bad elements of a movie with good ones. One example of such a movie is depicted in fig. 1 where the camera is shown both from the perspective of the prey and of the predator leading to several large lengths of the impending collisions as depicted in Fig. 2. During the climax, there is a direct contact, and therefore collision length is small and frequent. In this paper, we propose that contextual information can be used to aid our multimedia indexing system in recovering the semantic contents (high-level indices) of a video sequence. The popular video indexing systems (such as CHABOT [3], JACOB [4], and VisualSEEK [5]) employ the user composed queries in terms of the low-level features such as color, texture, shape etc., which are provided through the dialog box or by the way of selecting an example image. However, the process of information representation remains incomplete without the features, which are at a perceptual level like TTC. Features like TTC

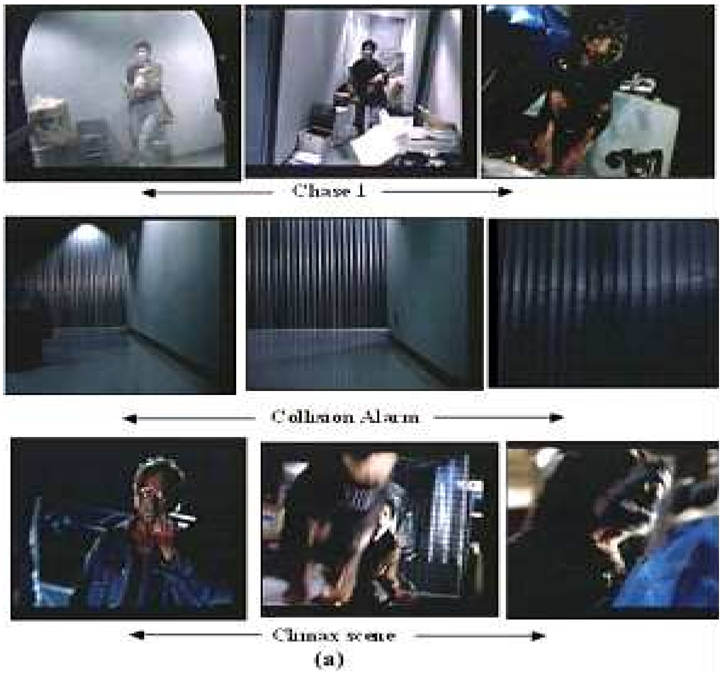


Fig.1: Shots leading to climax in a movie

reveal the fundamental structure about the content of the video data.

In this paper, we present a TTC algorithm to enhance characterization of video classes in higher-level categories. This will be realized in several steps. In Section II, we examine the conventional use of TTC shots in cinematography to affect the psychological state of the audience. This step is needed because the observations derived from this study can facilitate the formation of several TTC-based high-level indices from low-level features. In Section III, we attempt to extract a class of contextual information cues – *time-to-collision* (TTC) from the video sequence. The motivation behind this step stems from psychological studies which have shown that human beings are responsive to impending colliding objects [6], and from cinematographic studies which have shown that filmmakers often gradually shorten the camera-to-subject distance to intensify the audience's emotional involvement with the subject [7], to clarify detail [8], to identify objects of importance [9] or to amplify emotion [10]. Thirdly, in Section IV, we propose to refine our TTC detection algorithm using a simple and novel method, named as the *pilot cue*. Conclusions follow in section V.

2 Time-To-Collision (TTC) In Cinematography

In this section, the conventional use of TTC shots to affect the psychological state of the audience is presented.

2.1 Psychological Effects

Depending on how the cinematographer designed and juxtaposed the TTC shots in a scene, it can generate many special psychological effects.

2.1.1 Suspense

Suspense is generated when the audience is led to anticipate an exciting development or payoff. When creating a suspense scene, most directors tend to follow this cue-delay-fulfillment (payoff) pattern [11]. In general, camera movement may be used to create this delay effects. For instance, if the director slowly dollies the camera forward on a detail, gradually enlarging it but delaying the fulfillment of our expectations, the camera movement has contributed to suspense. In addition, if the director films a character moving slowly towards the camera, or if the director chooses to use a slow camera zoom in on a detail, the same delay effect will be created.

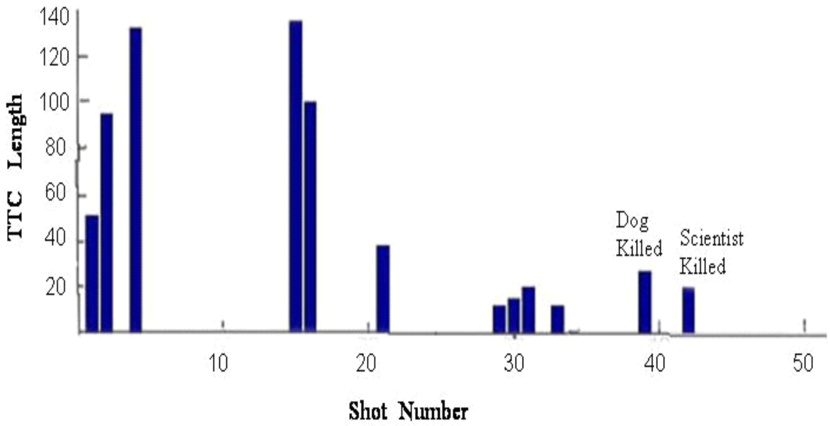


Fig.2: TTC values for the shots. Shots in which TTC length is not shown corresponds to case when TTC is infinite i.e., there is no possibility of collision.

2.1.2 Intimacy

By portraying the face in close up, the cinematographer makes it possible for the audience to know intimately the face of the character portrayed and hence to read his/her thoughts and feelings. As a result of that, the audience can get psychologically closer with the character. Moreover, by varying the camera-to-subject distance, the cinematographer can manipulate the audience's emotional involvement with the subject in complex ways. As the camera moves closer to a character, the audience is brought into that character's personal space which evokes greater intimacy between the audience and the character [12]. Hence, by zooming in or dollying towards (TTC) the subject's face or body, the cinematographer can call the audience's attention to significant facial expressions or gestures and thereby, allows the audience to get psychologically intimate with the subject.

2.1.3 Terror

Chosen for the lack of a better all-inclusive name, terror in our context can range from fear or panic to the shock portrayed by the characters in the scene. Resolution into each of these psychological states depends on what is being portrayed by the cinematographer. For instance, in a chase scene, the director can use a variety of camera movements to dynamically bring out the panicky state that the prey or victim is experiencing as the pursuer is chasing him/her (as shown in fig. 1). Such camera movements include dolly shots that offer the audience, views of the characters from the front as they run towards the camera, dolly shots that follow characters during the chase, and finally, dolly shots that offer the audience point of view of a certain character during the chase.

Conversely, the cinematographer can use camera zoom in on a character's face to intensify the fear or shock that the character is experiencing and also to facilitate the audience in interpreting the emotions experienced by the character. Since the zoom lenses are capable of magnifying the subject's face by many times, the audience will be able to see the facial expressions portrayed by the character more vividly and thus, the audience can read the feelings of the subject to a greater extent.

2.1.4 Others

Sometimes the cinematographer needs to direct the audience's attention to a particular object and to signify its importance to the audience. Zoom lenses are capable of performing this kind of attention grabbing technique since they can magnify the object of interest ten times or more, so that the audiences seem to move closer to that object of interest.

The dolly forward is always used to represent the movement of a character because the change in perspective caused by an actual change of position implies and indeed includes a real movement. If we see (in a long shot) someone sitting in the living room looking at a bottle of soft drink placed on the dinning table in the kitchen and if there is a dolly from his position up to the object, it would be surprisingly strange to see, in the course of the next shot, the man still sitting in the same place. Doubtless the audience will understand that the director is trying to represent the mental attitude of that person – an inclination of some kind [9]. Hence, the cinematographer can represent the mental inclination of a character to send messages or thoughts from his/her mind to another character through the use of dolly forward (TTC) shot.

2.2 Context-Based Interpretation

In cinematography, besides the special psychological effects that can be created through well-choreographed TTC shots in a scene, the cinematographer can also make use of contextual relations to affect TTC shots interpretation. Continuity in the events or the interpretations of the preceding or succeeding shots can affect the ways in which the audience perceives a TTC shot.

For instance, a zoom in (TTC) shot can be used within the context of an establishing shot to lead the audience into the environment in which the scene will occur. Moreover, the cinematographer can incorporate zoom in (TTC) feature into the context of other camera movements to direct the audience's attention to a significant object which is a key piece of story information.

3 Time-To-Collision (TTC) Detection

In this section, the methods used in this research to detect TTC shots in video sequence are explained and their mathematical formulations are presented. Finally, a discussion on the experimental results obtained is presented.

3.1 Movements causing TTC shots

Three kinds of movements in motion pictures that contribute to TTC shots are: (i) movement of the living beings and objects towards the camera; (ii) movement of the camera on a dolly towards a subject; and (iii) zoom in, an optical movement of the camera lens that apparently brings a subject closer to the audience from a fixed camera position.

3.1.1 Subject Movement – Moving Closer to the Camera

To emphasize the urgency of a character in escaping from a chase, the director tends to employ this class of TTC shot in which the character is running towards the camera. By filming a smooth and fast transition from a medium to a close up view of the character, the director seeks to gradually bring the character to be physically close to the viewers at this particular time of crisis and this greatly enhances the viewers' emotional involvement in the scene. An example of such shot is shown in Fig. 3 where Agent 007 tries to escape from his enemies after a fierce gunfight. In addition, the sense of urgency to escape from the enemies is further amplified through the use of wide-angle lens which makes the character's speed appear greatly accelerated. A special type of shot that falls under this class is known as the looming shot. In our context, looming refers to an optical stimulus arising from a symmetrical expansion of a closed contour in the field of view. These looming shots result in a visual experience, for a human observer, of an object on a collision course and when the closed contour of the object comes to fill the entire 1800 frontal view of view, a collision occurs. Hence, looming shot yields a three-dimensional perception for a human observer and this creates an illusion of depth for the viewer. An example of such shot (see Fig. 4) can be found in most action-packed movies where there is bound to be one or more explosion scene. In a typical explosion shot, a ball of fire (closed contour) will expand symmetrically from the point of destruction and this creates the illusion that the ball of fire is going to hit onto the face of the viewer.

3.1.2 Camera Movement – Dolly Forward

Dollying is actually a camera movement that causes the camera's viewpoint to change laterally or longitudinally. In this research, the main focus is on the camera movement towards a character or object and this movement is called dolly forward. When a camera dollies forward, the field of view covered by the camera becomes gradually narrow, thereby causing the background to be hidden progressively. Fig. 5 shows a sequence of images in which the camera is dollying towards two men standing in front of an airplane.

Cinematographers often employ this technique to represent the movement of a character, to identify objects of importance or to amplify emotion. Moreover, the dolly has the feel of independently moving through a space and in effect the dolly seems to bring the viewer into the scene. An interesting point to note is that as the camera dollies towards a character or object, the viewers will experience a change in perspective caused by the actual change in position of the camera and this change in perspective will not occur in the case of a zoom in.

3.1.3 Apparent Camera Movement – Zoom In

In this technique known as the zoom in, the camera remains in a fixed position and the zoom effect is created by continuously varying the lens focal length from normal to telephoto within a shot. This allows shots, like some camera movements, to go from very random information to very specific information. As the static

Fig. 3: A shot taken from **Tomorrow Never Dies** showing Agent 007 running along a corridor

Fig. 4: A shot taken from **Tomorrow Never Dies** showing the explosion of a mobile missile launcher

camera zooms in on a subject, the image will magnify but the perspective does not alter (contrast this phenomenon to the previous class of TTC shot i.e. dolly forward). This is because the camera does not move during the zoom in and hence the spatial relationships among objects remain constant, as a result of which the objects appear to be glued to their position and simply get bigger. Fig. 6 shows a zoom in shot of a terrorist arms trade bazaar.

Wherein the dolly forward seems to take the viewer into the scene, the zoom in seems to bring the scene to the viewer. Despite this difference, many directors tend to use the zoom like the dolly. It takes much less setup time and equipment, thus accomplishing roughly the same effect at a much cheaper price. In the faster working climate of television, it is particularly popular. As such, it has gotten a reputation as the poor man's dolly or the hack director's dolly. It is rarely used thoughtfully as a mean of representing different spatial approaches.

3.2 Implementation

A block diagram of our TTC detection algorithm is shown in Fig. 7. Our optical flow estimation algorithm is similar to the method proposed by Horn and Schunck [13], and Proesmans et al. [14]. The advantage of adopting these two algorithms is that a dense optical flow can be obtained, i.e. for every pixel in the image, an optical flow vector (u, v) can be obtained.

By fitting the affine model to the optical flow estimates obtained, the six affine parameters can be estimated using method of least squares as follows:

$$\begin{bmatrix} u_0 & v_0 \\ u_x & v_x \\ u_y & v_y \end{bmatrix} = \left[\sum_x \sum_y [1 \quad x \quad y]^T [1 \quad x \quad y] \right]^{-1} \times \sum_x \sum_y [1 \quad x \quad y]^T [u \quad v]$$

where $u_0, v_0, u_x, u_y, v_x, v_y$ are the affine parameters and x, y denotes the spatial coordinate.

The *time-to-collision* (TTC), which is the time that will elapse before the object and the camera collide, is computed as:

$$TTC = \frac{2}{div(\vec{v})} = \frac{2}{u_x + v_y}$$

where div is the divergence of the field. Here, we have made an assumption that there is no deformation present in the image flow field. The TTC is actually bounded by the deformation term (see [15]). A neutral approach was taken so that the presence of a strong deformation in the image flow field will not drastically affect the computation of the TTC.

After computing the TTC for each frame in the video sequence, the final step in our TTC detection algorithm is to recover the TTC shots from the video sequence. In

Fig. 5: A shot taken from **Tomorrow Never Dies** where the camera is dollying towards the two men standing in front of an airplane (from a car windscreen point-of-view)

Fig. 6: A shot taken from **Tomorrow Never Dies** where the camera is zooming in on a terrorist arms trade bazaar

a shot depicting an impending collision (i.e. having a large diverging image flow), the TTC values for every frame should be fairly small. If the TTC has a very large or negative value, then it simply means that the object is moving away from the camera and there is no possibility of a collision. After experimenting with some training video sequences, we set the threshold for the TTC value in each frame as 1000, i.e. for TTC value greater than 1000 or less than 0, a default high value of 10000 is set. Finally, the log of the average TTC value for each shot is then computed and evaluated.

3.3 Experimental Results

In this section, a video sequence taken from the action movie – *Tomorrow Never Dies* is used to evaluate the performance of our TTC detection algorithm. This video sequence was chosen because we desire to test the feasibility of extracting TTC shots from movies and also to investigate the problems that may arise from action-packed movies. The video sequence chosen has a frame rate of 25 frames/sec and is about 11 minutes long (total of 16,246 frames). The dimension of each image frame is 352 pixels by 288 pixels.

3.3.1 Performance Evaluation

The performance of the TTC detection algorithm is expressed in terms of recall and precision.

$$
\begin{aligned}
recall &= \frac{N_c}{N_c + N_m} \times 100\,\% \\
&= \frac{43}{43+14} \times 100\,\% = 75.4\,\%
\end{aligned}
$$

$$
\begin{aligned}
precision &= \frac{N_c}{N_c + N_f} \times 100\,\% \\
&= \frac{43}{43+16} \times 100\,\% = 72.9\,\%
\end{aligned}
$$

where *Nc*, *Nm* and *Nf* correspond to the number of correct detections, number of missed detections and number of false detections respectively. From the above statistics, we can see that the recall and precision for our TTC detection algorithm are well above seventy percent. Hence, our algorithm can be said to perform reasonably well under rather challenging data set.

3.3.2 Shot 1 – Correct Detection

Shot 1 is an image sequence whereby the object is moving closer to the camera. Fig. 8 depicts an example of

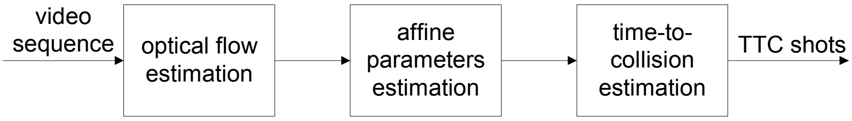


Fig. 7: Block diagram of the TTC detection algorithm

Frame 4 Frame 12 Frame 20

Fig. 8: The 4th, 12th and 20th frame of a shot showing a jeep moving closer to the camera

this type of shot that is correctly detected by our algorithm. In that shot, a jeep is moving closer towards the stationary camera and it will potentially collide with the camera if the cinematographer does not terminate the shot.

The TTC for each frame in shot 1 was computed and plotted as shown in Fig. 9. From the graph, we can observe that there are sharp jumps in TTC values. These sudden changes in TTC values are generally caused by the presence of independent moving objects in the image sequence. For instance, as shown in Fig. 8, we can observe that the man is running to the left while the jeep is moving forward. These two independent moving objects clearly violate our rigid scene assumption and hence, this violation leads to erroneous TTC being computed at certain frames in the shot. Fortunately, in our TTC detection algorithm, the average TTC value was taken for the entire shot and that helped to smoothen out any sudden fluctuation in TTC values.

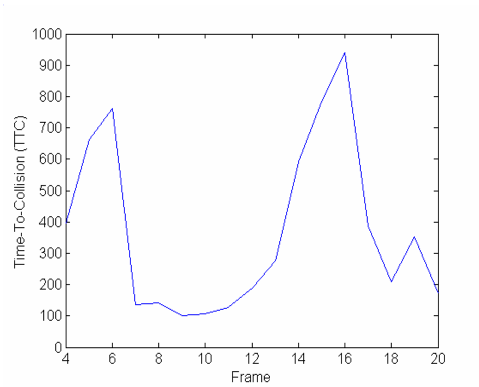


Fig. 9: Time-To-Collision (TTC) for each frame in shot 1

3.3.3 Lateral Movement – False Detection

Fig. 10 illustrates a lateral movement shot in which the camera is trying to track the movement of the man as he runs to the left. During the leftward tracking process of the camera, the background is actually moving to the right while the man is running in the direction of the camera movement and is running faster than the camera. Hence, there is an accelerated lateral flow in the horizontal direction (i.e. in the x direction) as depicted in the optical flow diagram in Fig. 10. We know that the divergence is directly proportional to the sum of u_x and v_y, where u_x and v_y are the rate of change of flow velocity (u, v) in the x and y directions respectively. Hence, the divergence computed for the above-mentioned shot will be small since there is an accelerated lateral movement in the x direction. This will cause our algorithm to misclassify this shot as TTC shot even though the man in Fig. 10 is not going to collide with the camera.

4 Improvement To TTC Detection – Pilot Cue

In airline pilots training school, pilots are taught that if another aircraft stays in the same location through their windscreen and grows large, they should immediately take evasive action as the aircraft that they see on their windscreen is on a direct collision course with their aircraft [16]. For instance, the aircraft that appears

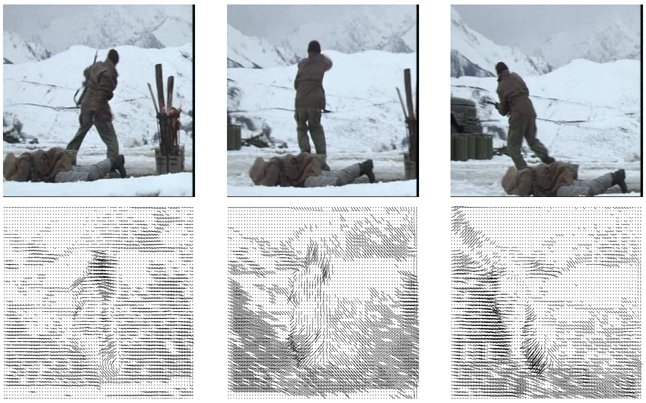


Fig. 10: The 4th, 13th and 22nd frame of a lateral movement shot and its optical flow – the shot is showing a camera trying to follow a man who is running to the left

on the pilot's windscreen, as shown in Fig. 11(a), will be on a collision course with the pilot's aircraft if sometime later, the aircraft stays in the same location on the windscreen and grows bigger in size, as illustrated in Fig. 11(b).

This method to detect potential collision can also be used in our TTC detection algorithm to extract TTC shots. For instance, if an object in the image frame stays in the same region throughout the whole shot and grows large then this shot will be classified as a TTC shot, since the object is on a direct collision course with the camera. Moreover, this method can also be used to filter out the false alarms that are due to lateral movement.

4.1 Method

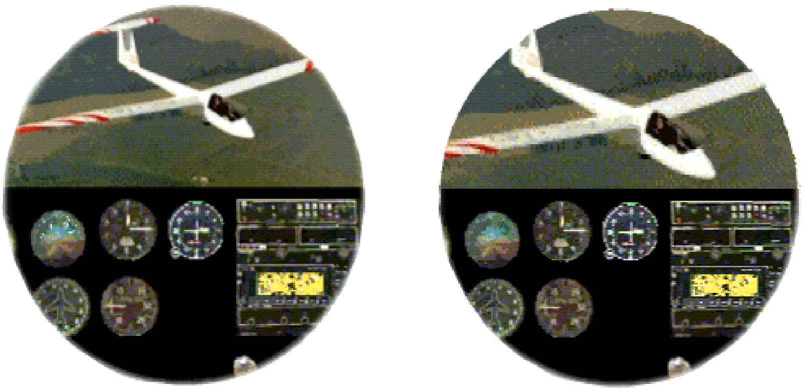
The approach that we take is to first divide the 352x288 image into 99 smaller square regions of size 30x30 (note that there is some offset at the boundary of the image). The TTC for each square region will be computed and a square will be labeled as TTC square if its TTC falls below a threshold of 100 (determined by some training sequences). Finally, the centroid of the TTC squares, as well as the total number of TTC squares, will be computed for each image frame. The centroids for the entire image sequence are said to stay in the same region of the image frame if the computed centroids fall within a circular boundary with center given by the center of the centroids and the radius given by 90 percent of the average flow strength for each pixel (percentage determined by some training sequences). The average flow strength is considered because it gives a rough indication of the amount of spatial movement present in the image. As for the number of TTC squares, we plot these numbers along the time axis and use least square fit to find the best fitting straight line through these points. A positive line gradient will indicate that the number of TTC squares is increasing. If the image sequence satisfies all the above-mentioned conditions, it will be labeled as a TTC shot.

4.2 Experimental Result – Sequence 1

Sequence 1 is a 38 frames video sequence whereby a man is running towards the camera. Fig. 12 illustrates the 4th, 19th and 35th frames of the video sequence and their corresponding optical flow diagrams. The centroid of the TTC squares for each frame in the video sequence is computed and plotted as shown in Fig. 13(a). All the centroids in this video sequence were found to lie within the circular boundary with center given by the center of the centroids – (7.08, 4.22) and radius given by 90 percent of the average flow strength for each pixel – 2.55. Hence, the first condition that the centroids must stay within the same region in the image frame is satisfied.

The number of TTC squares for each frame in the video sequence is also plotted as shown in fig 13(b). Using least square fit, the gradient of the straight line that best fits the curve was found to be 0.39 squares/frame. Hence, the second condition that the number of TTC squares must increase temporally is also satisfied.

Clearly from the analysis conducted, the *pilot cue*, which requires the object to stay in the same region of



(a) (b)

Fig. 11: (a) A view of an aircraft on a windscreen and (b) the aircraft has stayed in the same location and has grown bigger in size

the image frame and grows large, has been satisfied. Hence, we have shown that it is feasible to use *pilot cue* to detect TTC shot where the object is moving towards the camera and there is impending collision.

4.3 Sequence 2 – Non-Collision

Sequence 2 is a 57 frames documentary video sequence whereby a wild boar is running towards and across the viewing camera. The image size for this video sequence is 352x240 instead of the usual 352x288 and hence, the image is divided into 77 smaller square regions each of size 30x30 (offset at the boundary is still provided). Fig. 14 shows the 4th, 29th and 54th frames of the video sequence and their corresponding optical flow diagrams. Visually, it is easy to notice that although the wild boar is running closer to the camera, it is not on a direct collision course with the camera. Hence, the goal here is to show that this video sequence does not satisfy our pilot cue – i.e. although the wild boar in the video sequence is growing bigger in size, it is not staying within the same region in the image frame.

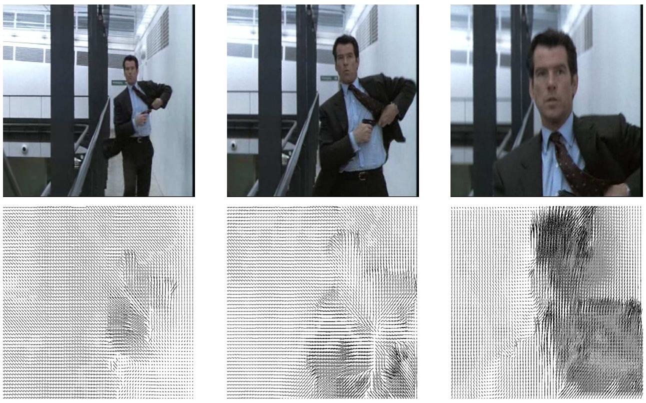
Using the method as described earlier, the centroid of the TTC squares for each frame in the video sequence is computed and plotted as shown in Fig. 15(a). Most of the centroids are found to lie in the lower half of the image frame which corresponds to the region of the wild boar's movement. Also, the centroids do not lie within the circular boundary with center defined by the center of the centroids – (6.14, 3.89) and radius defined by 90 percent of the average flow strength – 1.10. Hence, the first condition that the centroids must stay within the same region in the image frame is not satisfied. This implies that the wild boar is not staying in the same location on the image frame.

Next, the number of TTC squares is computed for each of the frame in the video sequence and plotted as shown in Fig. 15(b). From the figure, we can observe that the gradient of the straight line obtained using least square fit is positive and is found to be 0.13 squares/frame. Hence, the second condition of the pilot cue is satisfied, i.e. the total number of TTC squares is increasing. This implies that the wild boar is moving closer to the camera. However, there is no impending collision with the camera since the video sequence does not satisfy the first condition of the pilot cue. From the above analysis, we have shown that it is also feasible to use pilot cue to detect an object that is moving closer to the camera but is not on a direct collision course with the camera. Hence, the pilot cue will be extremely useful in our TTC detection algorithm to reduce the number of false detections that are caused by lateral movement. We just need to show that the centroid of the object undergoing lateral movement does not stay in the same location throughout the whole shot.

4.4 Discussion

The module TTC contains the following sub-descriptors for linking to high-level video categories:

1. Duration of impending collision, i.e., the number of frames for which possible collision is detected.



Frame 4 Frame 19 Frame 35

Fig. 12: The 4th, 19th and 35th frames of a video sequence and their corresponding optical flow – the video sequence is showing a man running towards the camera

This length varies depending on the type of media. For action movie or sports, it is small (sudden collision), while for climax or chase sequences (like in an animal documentary) it is large, which creates empathy of apprehension with the prey.

2. Frequency of occurrence of impending collisions, i.e., the number of times low TTC value is detected in a fixed time. For media which is mainly a shot from a medium-distance or a short-distance camera, the frequency of impending collision occurring would be high. For example, in a rugby video or in a violence video, there tends to be a large number of collisions detected over a short time.

5 Conclusions

In contrast to the physical characteristics such as color, texture, optic flow etc., perceptual features such as time-to-collision form much powerful cues in the characterization of many video classes. Temporal fusion of these features form higher-level structures, which can be effectively formulated to characterize emotion which is not possible with low-level features. The goal of this paper is to propose an algorithm for accurate detection of TTC in movie databases and to explore its use in detecting semantics in a video. In future, more sophisticated approaches could be used to refine rule-based approach (for example, using Dynamic Bayesian Networks [17]). In addition, to achieve the complete framework, future work may be done to include other features pertaining to high-level indices such as camera movement, subject movement and audio input.

References

1 F. G. Meyer, "Time-to-collision from First-order models of the motion fields", *IEEE Transactions of robotics and automation,* pp. 792-798, 1994

2 D. Marr, Vision: A computational investigation into the human representation and processing of visual information, Freeman. 1982

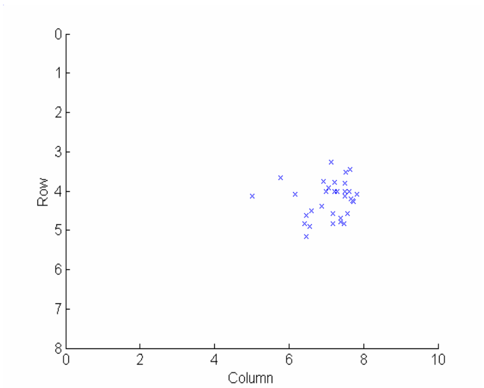
3 V. E. Ogle and M. Stonebraker. CHABOT: Retrieval from a relational database of images. *IEEE Computer*, pages 40–48, September 1995.

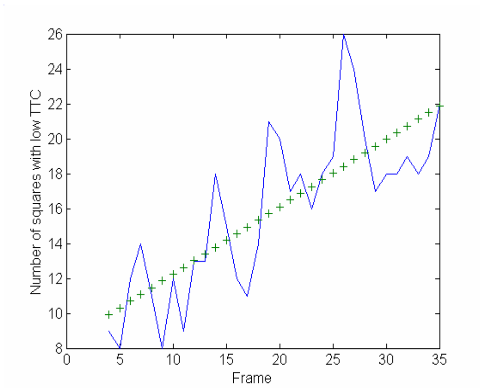
4 M. L. Cascia and E. Ardizzone. JACOB: Just a contentbased query system for video databases. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pages 1216–1219 May 1996.

5 J.R. Smith and S. F. Chang. VisualSEEK: a fully automated contentbased image query system. *ACM Multimedia*, November 1996, pages 87–98.

6 W. Ball and E. Tronick, "Infant Responses to Impending Collision: Optical and Real", *Science,* New Series, Vol. 171, Issue 3973, pp. 818 – 820, Feb 1971

7 S. Prince, "Movies and Meaning: An Introduction to Film", Allyn & Bacon, 1997





(a) (b)

Fig. 13: (a) A scatter plot for the centroid of the TTC squares in each frame of video sequence 1 (b) A graph of the number of squares with low TTC against the frame number for video sequence 1

8 J. M. Boggs, “The Art of Watching Films”, Fourth Edition, Mayfield Publishing Company, 1996

9 J. Mitry, “The Aesthetics and Psychology of the Cinema translated by Christopher King”, Indiana University Press, 1998

10 B. Mamer, “Film Production Technique: Creating the Accomplished Image”, Second Edition, Wadsworth – a division of Thomson Learning, 1999

11 W. Miller, “Screenwriting for Film and Television”, Allyn and Bacon, 1998

12 R. B. Peacock, “The Art of Movie Making: Script to Screen”, Prentice Hall, 2001

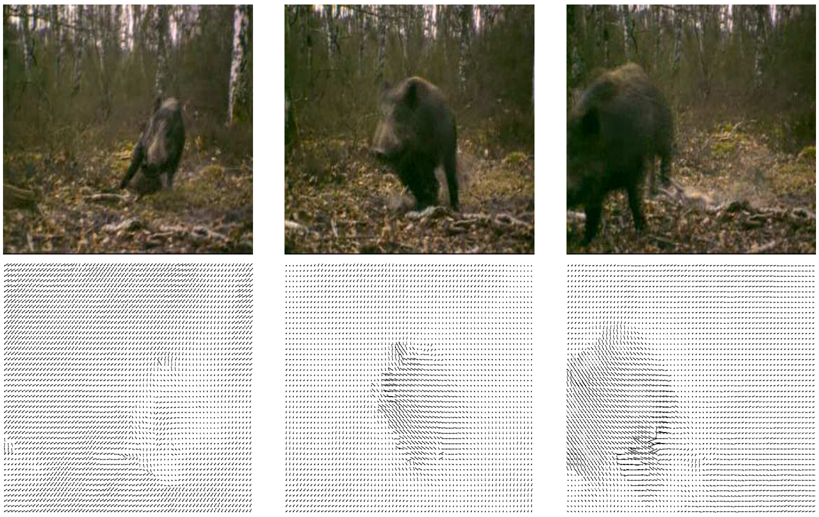
13 B. K. P. Horn and B. G. Schunck, “Determining Optical Flow”, *A.I. Memo No. 572*, April 1980

14 M. Proesmans, L. Van Gool, E. Pauwels and A. Oosterlinck, “Determining of Optical Flow and its Discontinuities using Non-Linear Diffusion”, *3rd European Conference on Computer Vision*, Vol. 2, pp. 295 – 304, 1994

15 M. Subbarao, “Interpretation of Visual Motion: A Computational Study”, Pitman Publishing Limited, 1988

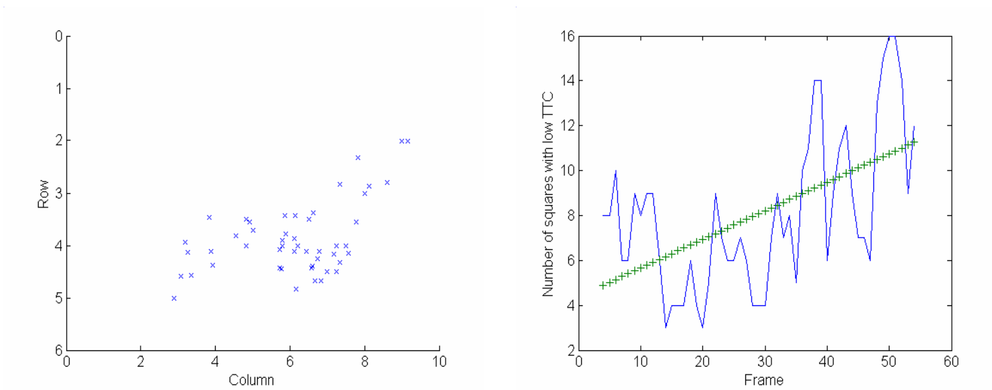
16 J. E. Cutting, P. M. Vishton and P. A. Braren, “How We Avoid Collisions With Stationary and Moving Obstacles”, *Psychological Review*, Vol. 102, No. 4, pp. 627 – 651, 1995

17 A. Mittal, L.-F. Cheong and T.-S. Leung, “Dynamic bayesian framework for extracting temporal structure in video", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’01)*, Hawaii, USA December 2001, Vol. II, pages 110-115



Frame 4 Frame 29 Frame 54

Fig. 14: The 4th, 29th and 54th frames of a video sequence and their corresponding optical flow – the video sequence is showing a wild boar running across the viewing camera and there is no impending collision



(a) (b)

Fig. 15: (a) A scatter plot for the centroid of the TTC squares in each frame of video sequence 2 (b) A graph of the number of squares with low TTC against the frame number for video sequence 2