# Various Camera Motion Type Estimation of Animation Sequences

Hailong Jiang, Guoxu Liu, Jae Ho Kim PhD

Department of Electronic Engineering, Pusan National University

## INTRODUCTION

A large body of work in computer vision over the last several decades has been concerned with the extraction of motion information from image sequences. The purpose of this work is actually quite diverse, such as video compression, video indexing, structure from motion, visual story telling and object tracking. Moreover, with the development of multimedia market, the need for the applications above becomes more and more so as to the role of camera motion estimation (CME) is bigger than ever before. However, up to now, most of the CME methods have focused on the 6 types of camera motion, and they are far from being satisfactory for various applications. In this research, a new scheme of CME which can estimate up to 18 camera motion types is proposed.

## OBJECTIVES

In this study, the research is focused on estimating 18 CM types with considerate efficiency and eliminating the problem of rapidly moving character (object). The objectives of this research are listed as follows.

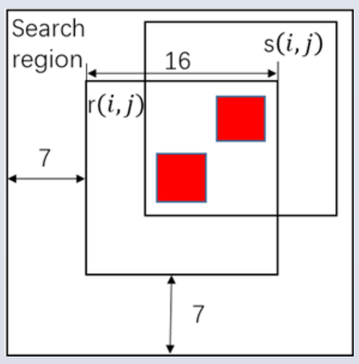
- Estimation of 18 camera motion types with considerable efficiency.
- Elimination of the problem of rapidly moving character (object).

## MATERIALS AND METHODS

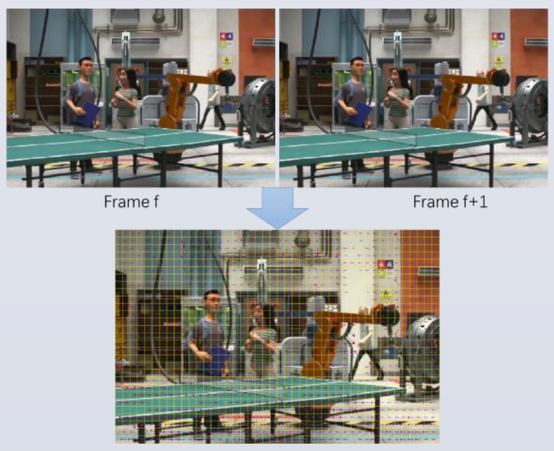
The proposed method consists of several steps and are presented as follows.

### Block Matching

The block matching algorithm is a method of finding the most similar blocks between consecutive frames. In this research, it is adopted to find the motion vectors for camera motion estimation.




The macroblock and search area



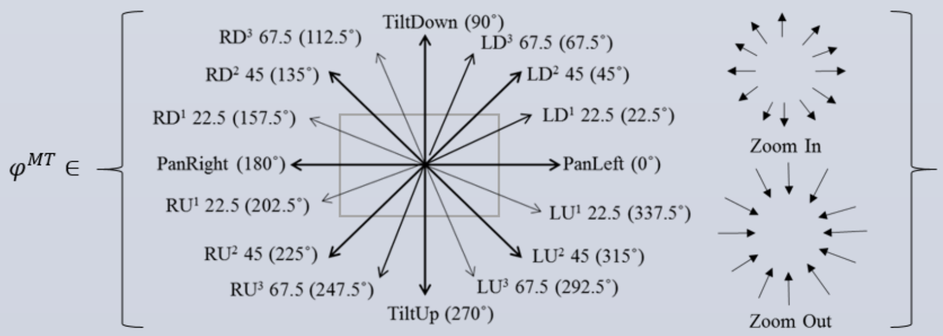

The motion vectors generated by BMA

### Global Motion Filtering

Y. F. Ma proposed the global motion filter[1] to extract actual object motion from the mixed motion of object and camera. This research applies the global motion filter to distinguish background and objects of a video clip.

### Motion Estimation using Minimum Sum of Squared Difference

In this study, 18 types of camera motions can be estimated and the ideal motion vector angles for these camera motions are also defined.

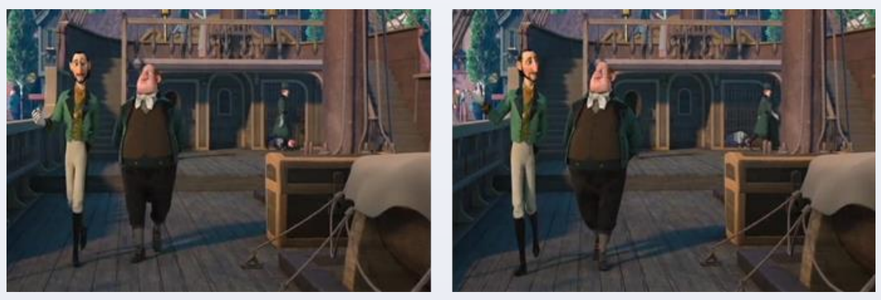



Types of camera motions to be distinguished in this research

## RESULTS AND DISCUSSIONS

Movie animations are used as experimental data to test the proposed method. In this research, 'Frozen' (2013) and 'Big Hero' (2014) which got the 71th and 72th Golden Globe Awards, respectively, are selected for experiment.

Using the proposed method, we get a good result for the experiment. One example of the result is shown in the following figure.



An example of good result in the experiment (PanLeft)

| | *Pan-right* | *Pan-Left* | *Tilt-Up* | *Tilt-Down* | *RU*[1] | *LU*[1] | *RD*[1] | *LD*[1] | *RU*[2] |
|---|---|---|---|---|---|---|---|---|---|
| $E^{MT}$ | 1.35E+05 | **6.90E+03** | 7.27E+04 | 6.66E+04 | 1.25E+05 | 2.18E+04 | 1.17E+05 | 1.91E+04 | 1.07E+05 |
| | *LU*[2] | *RD*[2] | *LD*[2] | *RU*[3] | *LU*[3] | *RD*[3] | *LD*[3] | *Zoom-In* | *Zoom-Out* |
| $E^{MT}$ | 3.86E+04 | 1.00E+05 | 3.44E+04 | 9.00E+04 | 5.56E+04 | 8.33E+04 | 5.03E+04 | 7.57E+04 | 6.50E+04 |

The estimated camera motion type

To evaluate the effectiveness of this research, the proposed method is compared with S.G. Ma's method.[2] Except zoom, the proposed method distinguishes 16 translational CMs. But S.G. Ma's method distinguishes only 4 (pan left/right and tilt up/down). Therefore, the average of 16 CMs in the proposed and average of 4 CMs in Ma's methods are compared.

| | **Precision** | | **Recall** | | **Accuracy** | |
|---|---|---|---|---|---|---|
| *Method* / *CMT* | *The Proposed* | *Ma's Method* | *The Proposed* | *Ma's Method* | *The Proposed* | *Ma's Method* |
| ***Average of 16 CMs*** | 91.56% | - | 92.59% | - | 99.03% | - |
| ***Average of Pan and Tilt*** | - | 83.2% | - | 83.7% | - | 92.4% |
| ***Zoom*** | 92.05% | 86.9% | 86.1% | 78.4% | 97.3% | 96.7% |

Comparison with S. G. Ma's method.

## CONCLUSIONS

In this research, a new scheme of CME adopting the minimum sum of squared difference of the block motion vector angles is proposed. Ideal motion vectors of the blocks are suggested for 18 CM types. The Global Motion Filter is employed for removing the moving character (object) of the animation video clips. The proposed method shows 91.56% precision, 92.59% recall, and 99.03% accuracy for the average of 16 translational CMs. These are 8.36%, 8.89%, and 6.63% higher in precision, recall, and accuracy, respectively compared to S. G. Ma's method. The proposed method shows 92.05% precision, 86.1% recall, and 97.3% accuracy for zoom. These are 5.15%, 7.7%, and 0.6% higher compared to the zoom of S.G. Ma's method.

## REFERENCES

1. Y. F. Ma & H. J. Zhang "A new perceived motion based shot content representation," in *Proc. IEEE Int. Conf. on Image Process. (ICIP)*, **3**, pp. 426-429 (2001).
2. S.G. Ma & W.Q. Wang, "Effective camera motion analysis approach," in *IEEE International Conf. on Networking, Sensing and Control (ICNSC'10)*, pp. 111-116 (2010).

## CONTACTS

Pusan National University
2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, 46241, Rep. of Korea
Email : jhkim@pusan.ac.kr