A Surveillance System in the Skiing Ground

Koji UENO and Masahiro IWAHASHI
Department of Electrical Engineering, Faculty of Engineering
Nagaoka University of Technology
kueno@tech.nagaokaut.ac.jp

Abstract - This paper introduces a surveillance system in the skiing ground which can automatically extract individual player on the snow-covered slope. Robustness of the extraction against various weather conditions such as sunshine, snowing, raining, etc. are discussed as main technical problems. The system is now under development so that it will be utilized for detecting accidents and monitoring the area for safety and amenity.

I. INTRODUCTION

So far, numerous surveillance systems with digital video camera have been developed for various applications such as tracking cars, people, players on the road, pedestrian street, soccer field, respectively [1-2].

This paper introduces experimental results of our surveillance system on the "skiing ground" and describes future plan of our research. Robustness of the extraction against various weather conditions such as sunshine, snowing, raining, etc. are discussed as main technical problems. Dynamic background update, background subtraction, labeling, clustering and matching techniques are utilized to extract and track an individual player on the snow-covered slope. Accuracy of the extraction is improved by considering some conditions peculiar to the "skiing ground".

The system is now under development so that it will be utilized for detecting accidents and monitoring the area for safety and amenity.

Architecture of the system and basic video signal processing techniques are summarized in section II. Our devices for "skiing ground" are proposed in section III. Experimental results are introduced for evaluation of the system in section IV. Conclusions are stated in section V.

Table 1 What the system intends to detect.

Objective ↑	Position of a player. Velocity, Acceleration. Number of the players.
	Intrusion into a forbidden area. Toppling, Slipping. Collision, Crushing. Density of the players. Skiers or boarders or snowmobile. Tendency or behavior as a group.
↓ subjective	Safety and Amenity.

II. VIDEO SIGNAL PROCESSING

A. Architecture of the System

The system is intended to detect the items listed in **table 1** with a "single" video camera fixed on the ground or floor in a room. Sensor fusion is out of the range in this report. As for the detection, we will face various problems as below.

- Sunshine change (Influence of moving clouds)
- Subtle change of the background (moving objects by wind or something)
- Weather condition (snowing, raining, fog)
- Color of the player similar to the background (difficult to distinguish from snow)
- Occlusion (Separate two players close to each other)

Basic algorithm of the video processing to extract an individual player in the system is illustrated in **figure 1**. The "dynamic background update" makes a subtle change of the background negligible. The "background subtraction" extracts moving areas in the scene. The "labeling and clustering" identify a player connecting some moving areas. The "matching and tracking" trace where the player and make the system robust against the occlusion of players.

Each of the basic tools are summarized below and special turnings to the skiing ground are proposed in section III.

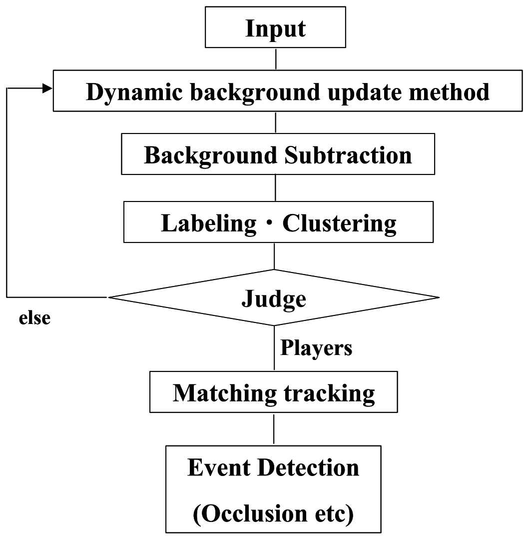


Figure 1 Basic video signal processing algorithm.

B. Dynamic Background Update

Main purpose of the system is to extract players (foreground) from the scene (background) in video signal. Under the gradually changing sunshine due to direction of the light and moving clouds, the background also gradually changes. Therefore, it is necessary to renew the background before the "background subtraction" described in the next subsection.

Denoting $B_t(x,y)$ and $I_t(x,y)$ as a luminance value of the background and the captured frame at time *t* and location *(x,y)* respectively, the "Dynamic Background Update" [3] is defined by

$$B_t(x,y)=\begin{cases} B_{t-1}(x,y) & \cdot\ \cdot\ k<A_o(x,y) \\ \dfrac{I_t(x,y)+(n-1)*B_{t-1}(x,y)}{n} & \cdot\ \cdot\ k>A_o(x,y) \end{cases} \quad (1)$$

where *k* is a threshold calculated with the least mean square method and A_0 is a value from the "background subtraction" in the next subsection.

Advantage of this method is that it is not necessary to store a few frame data and it requires only one calculation for one pixel. Therefore the method is memory and time saving. However, when a sudden light change occurs (e.g. when a thick cloud suddenly covers the sunshine), this method does not work well and it is impossible to renew the background. This problem is dealt with in the section III.

C. Background Subtraction

The "background subtraction" extracts a moving area by taking subtraction between a captured frame A_i and the background area in the previous frame A_b by

$$A_o(x,y)=\left|A_b(x,y)-A_i(x,y)\right| . \quad (2)$$

If the threshold *k* in the previous subsection is a fixed value, the background and foreground separation can not be adaptive to natural change of light. Therefore we calculate the threshold *k* with the least mean square method by

$$\varepsilon^2(k^*)=\min_{0<k<255}\varepsilon^2(k)$$
$$\varepsilon^2(k)=\sum_{i=0}^{k}(i-\mu_1)^2 p_i+\sum_{i=k+1}^{255}(i-\mu_2)^2 p_i \quad (3)$$
$$\mu_1=\frac{\sum_{i=0}^{k}p_i\cdot i}{\sum_{i=0}^{k}p_i}\ ,\quad \mu_2=\frac{\sum_{i=k+1}^{255}p_i\cdot i}{\sum_{i=k+1}^{255}p_i}$$

where p_i denotes a probability density function of a pixel whose luminance value is equal to *i*. As a result, it becomes possible to automatically and also dynamically separate the background and the foreground against various noise such as light change.

D. Labeling, Clustering and Player Extraction

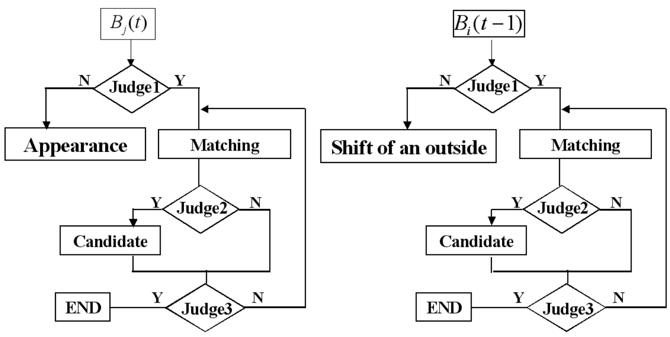
The "background subtraction" normally gives us a few pixel groups corresponding to a single person. The labeling put the same label on the pixels which share almost the same luminance value. The clustering makes some pixel groups each of which is composed of neighboring pixels under the same label. As a result, we can expect to have pixel groups (blobs) each of that corresponds to one player.

However the algorithm requires us further improvement to achieve the perfect extraction of all the players. It is actually impossible but our efforts are introduced in section III.

E. Matching and Tracking

Tracking technique of the players is necessary to detect collision (clash) or occlusion of a few players. The tracking in this report takes "matching" between the blob in the current frame and that in the previous frame. Algorithm of this technique [2] is illustrated in **figure 2**.

It becomes possible to recognize whether the player disappears out of or appears into the scene. Our efforts to improve accuracy of tracking players in the skiing ground is introduced in section III.



(a) Current blob. (b) Former blob.

--- Judge 1 ---
Is there any adjacent blob in the previously detected blobs?
--- Judge 2 ---
Is the blob optimum among the detected blobs?
--- Judge 3 ---
Is there any other neighboring blobs?

--- Decision 1 ---
Is there adjacent blob in the current blobs?
--- Decision 2 ---
Is the blob optimum among the detected blobs?
--- Decision 3 ---
Is there any other neighboring blobs?

Figure 2 A player tracking algorithm.

III. ADAPTATION TO THE SKIING GROUND

A. Player's Location and Square Measure

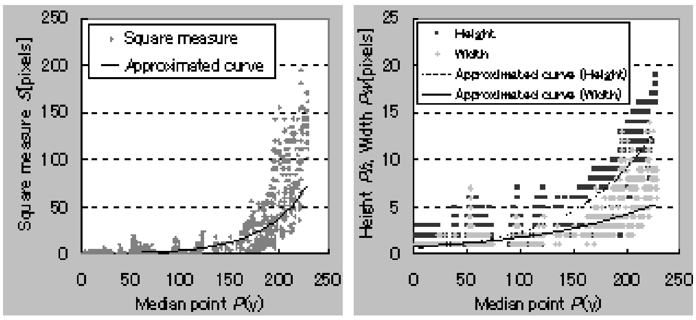
To improve accuracy of the recognition, an assumption of the system is utilized. A camera set to the ground at the bottom of the valley looks up to the hill. In this situation, there is a correlation between location P_y (center of the blob), height P_h, width P_w and square measure S of a player. Namely, the closer to the camera a player is , the lager in the video its size becomes. **Figure 3** indicates experimentally measured results for a skiing ground. Recurred curves are also illustrated.

Our system employs these curves. Only the blobs close to these curves are recognized as the blobs related to a player. As a result, some erroneous blobs due to a sudden light change by a moving cloud which covers the sunshine are excluded from recognition process. In **figure 4**, we can see white colored wide area at the top of the hill. Such area is excluded since its squared measure is too big to be a player. Similarly, too small blobs are also excluded.

B. Improvement of Matching and Tracking

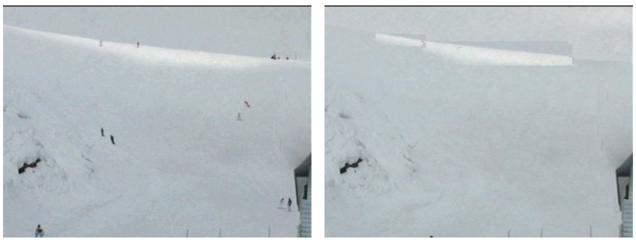
Another assumption in which upper body of a player is more stable than that of lower body is utilized for more accurate recognition. **Table 2** compares matching with upper, lower and whole part of the blob, respectively. It indicates difference between two blobs in the current frame and in the previous frame in the matching procedure. Values are calculated as average of 100 samples.

As a result, we can conclude that matching with "upper" part of the blob is recommended. In this case, processing speed is increased since size of the blob is reduced.



(a) S and $P(y)$ (b) P_h, P_w and $P(y)$

Figure 3 Relation between S, P_h, P_w and $P(y)$ of the objects.



(a) Original image. (b) Background image.

Figure 4 Background update for sunshine change.

IV. EXPERIMENTAL RESULT

A. Accuracy of Extraction and Tracking

The experiments in this report are performed under the configuration summarized in **table 3**. Extraction ratio under various weather condition, fine, cloudy and partly cloudy, are indicated in **table 4**. If the recognition is perfect (without any erroneous recognition) then the figure becomes 100%. Under the cloudy weather, the ratio decreases since color of the player is not bright and it becomes difficult to distinguish it from the snow in the background.

Tracking rate which indicates accurate tracking of the players is summarized in **table 5**. As we have expected, the matching with upper part of the blob is better that that with whole blob under various weather.

B. Examples of Extraction, Matching and Tracking

An example of the extraction of a player is illustrated in **figure 5**. Squared area with red line is the blob recognized by the system. It becomes difficult when the players are at a long distance from the camera since only one or more pixels are assigned to a player. We have also confirmed that the recognition is robust against snowing condition since the snow has one or less pixels in the video data. However, if it is impossible for us to see the players, then it is also impossible for the machine!

An example of the tracking is illustrated in **figure 6**. Trace of the center of the players are indicated as green lines. This procedure requires memory to store previous location of the blobs, however, it is expected to find status of the skiing ground such as velocity and acceleration of the players.

Table 2 : Matching results.

	Body parts		
	Upper	Lower	Whole
Skier	9.29	14.07	11.68
Snowboarder	4.20	6.28	5.54

Table 3 : Sequence characteristic.

Sequence time	60[sec]
File type	AVI
Process unit	Bitmap (256[tones])
SIZE	320×240[pixels]
framerate	25[frames/sec]

Table 4 : Extraction ratio.

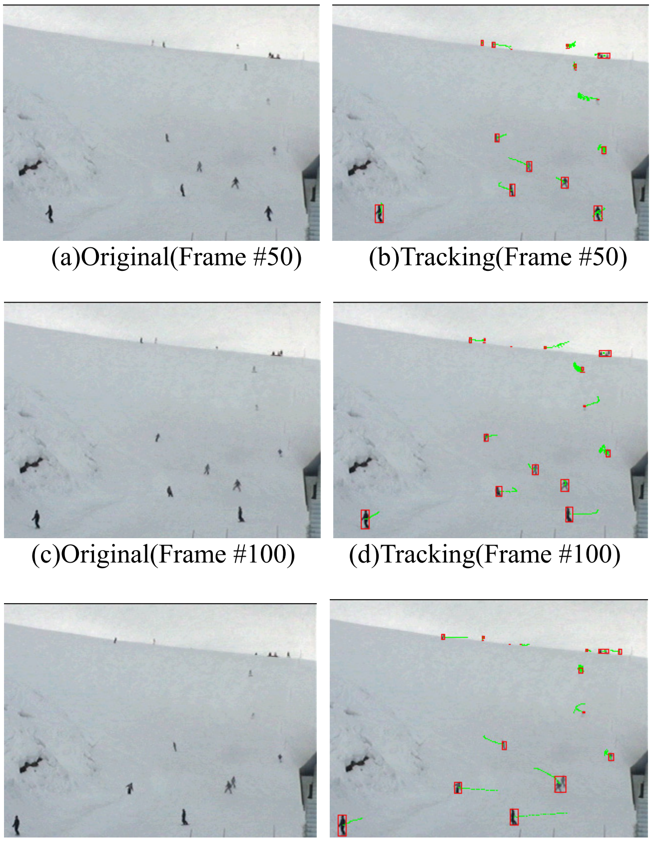
Test sequence	extraction ratio[%]
fine	94.02
cloud	91.53
fine and cloud	93.01

Table 5 : Tracking ratio.

Test sequences	Player part	
	whole[%]	Upper[%]
fine	90.26	92.38
cloud	89.58	91.72
fine and cloud	88.16	91.55

(a) Original image. (b) Extracted image.

Figure 5 Examples of the player extraction.

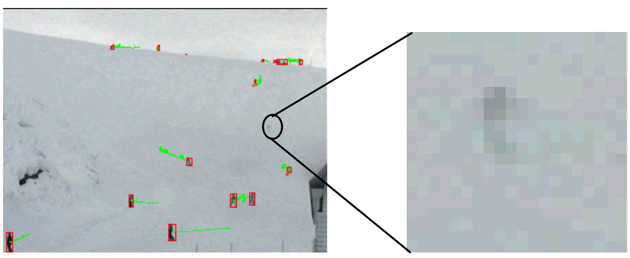


(a)Original(Frame #50) (b)Tracking(Frame #50)

(c)Original(Frame #100) (d)Tracking(Frame #100)

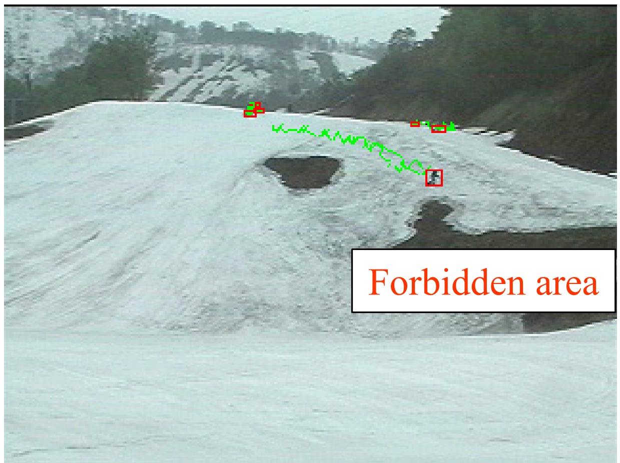
(e)Original(Frame #150) (f)Tracking(Frame #150)

Figure 6 Examples of the player tracking.



(a) Tracking image.(#165) (b) Enlarged image.

Figure 7 Example of extraction failure.



(a) Intrusion in forbidden areas.

(b) Occlusion of two players.

Figure 8 Examples of event detection.

Table6: Processing speed.

	Processing speed[msec]
Matching (half)	4.396
Matching (whole)	10.371
System	15.891

Table7 : Event detection examples.

Event detection	Metadata
People Counting	blob ID, 2-D pos.
Intrusion	blob ID, 2-D pos.
Slip	dy/dx, S

C. Examples of Failures and Event Detection

An example of failure of the extraction is indicated in **figure 7**. In this case, the area of the player is almost same as the background.

Some events such as intrusion into a forbidden area can be detected as indicated in **figure 8(a)**. Collision or crushing can be detected as illustrated in **figure 8(b)**. However, in this case, we do not know whether it is a serious matter which needs rescue or not. At this moment, this system can not perfectly detect an occlusion of two players, it sometimes recognizes this situation as a single big player!

D. Processing Time

Measured signal processing time is summarized in **table 6**. The system takes 16 [msec] per one frame and it is feasible to implement a real time recognition system.

V. CONCLUSIONS

A surveillance system in the skiing ground which can automatically extract individual player on the snow-covered slope was introduced. Robustness of the extraction against various weather conditions such as sunshine, snowing, raining, etc. are increased by utilizing assumption on the skiing ground and players. Real time implementation is realized for the system.

The system is now under development so that it will be implemented for detecting accidents and monitoring the area for safety and amenity. In the future, toppling, slipping, density of the players will be detected. It should be categorized whether skier, snow boarder or snowmobile. Tendency or behavior as a group will be analyzed.

REFERENCES

[1] L.M.Fuentes, S.A.Velastin, “From Tracking to Advanced Surveillance”, IEEE International Conference on Image Processing (ICIP), 2003.

[2] L.M.Fuentes, S.A.Velastin, “People Tracking in Surveillance Applications”, 2nd IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), 2001.

[3] R.Shimada, K.Matsuda, T.Yagishita, H.Ushijima, T.Mori, H.Yamada, “A Dynamic Background Image Extraction Method for the Inside Invader Detection”, Technical Report of IEICE. IE2001-45, PRMU2001-65, 2001.

[4] N.Otsu: "A Threshold Selection Method from Gray-Level Histograms," IEEE Trans. Syst. Man & Cybern., Vol.SMC-9, No.1, pp.62--66, 1979.