2013 Second IAPR Asian Conference on Pattern Recognition

Group Leadership Estimation Based on Influence of Pointing Actions

Hitoshi Habe
Dept. of Informatics
School of Science and Engineering
Kinki University
Osaka, JAPAN 577-8502
Email: habe@computer.org

Kohei Kajiwara Ikuhisa Mitsugami Yasushi Yagi
Dept of Intelligent Media
Institute of Scientific and Industrial Research
Osaka University
Osaka, JAPAN 567-0047

***Abstract*—When we act in a group with family members, friends, colleagues, each group member often play the respective role to achieve a goal that all group members have in common. This paper focuses on leadership among various kinds of roles observed in a social group and proposes a method to estimate a leader based on an interaction analysis. In order to estimate a leader in a group, we extract pointing actions of each person and measure how other people change their actions triggered by the pointing actions, i.e. how much influence the pointing actions have. When we can see the tendency that one specific person makes pointing actions and the actions have a high influence on another member, it is very likely that the person is a leader in a group. The proposed method is based on this intuition and measures the influence of pointing actions using their motion trajectories. We demonstrate that the proposed method has a potential for estimating the leadership through a comparison between the computed influence measures and subjective evaluations using some actual videos taken in a science museum.**

I. Introduction

In order to provide an appropriate service for guests in a public space such as museums and department stores, what kinds of people they are and what they are doing are quite important information. For example, the information suitable for families would be different from one for office workers. Moreover, the way by which information is provided should be changed according to the type of the guests. Based on this motivation, many research groups have been studying how to estimate the attributes of people observed in a video. This would be a key technology for a computer system to provide fruitful information for guests and to assist their daily activities in a casual manner.

This work is also based on this motivation but tackling a more challenging task. We focus on collective activities and respective roles played by each member of a group. When we act in a group with family members, friends, colleagues, each group member often play the respective role to achieve a goal that all group members have in common. If it is possible to estimate roles of people in a group, the more appropriate information can be provided.

Among various kinds of roles in group activities this paper focuses on leadership. Leadership can be observed in many situations. For example, when a group is talking with each other to decide what they will do next, we often observe that a particular person is leading the conversation. After deciding a following action, he/she is likely to continue leading the group by instructing other members.

The ultimate goal of our work is to estimate a leader of a group. That enables us to find out a 'key person' of a group for advertising a product as illustrated in Fig. 1. Needless to say, as shown in the examples above, the term 'leader' contains numerous aspects. It is not reasonable to aim at estimating a leader exhaustively, covering such numerous aspects. We should break down the problem into a feasible one.

A leader-follower relationship can be visually observed via an instruction between a leader and a follower as shown in Fig. 2. This would be a reasonable cue as a first step toward our final goal because instructions can be detected using computer vision techniques.

Our proposed method extracts pointing actions of each person and measure how other people change their actions triggered by the pointing action, i.e. how much influence the instruction corresponding to the pointing action has. When we can see the tendency that one specific person makes pointing actions, i.e. giving instructions and the actions have a high influence on another member, it is very likely that the person is a leader in a group.

Some works that treat group activities have been already shown in the literature. Collective activity recognition is one of the emerging research topics among computer vision community [1]–[4]. Besides, interaction between people can be used

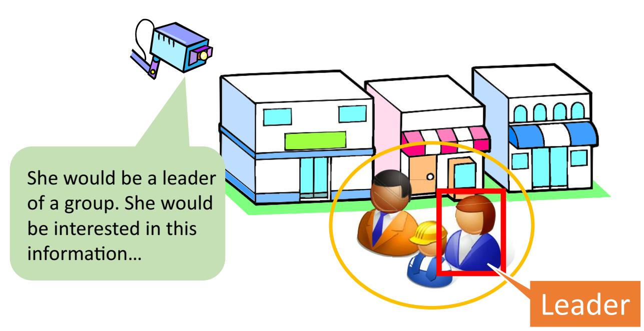


Fig. 1. Application Scenario of Leadership Estimation

978-1-4799-2190-4/13 $26.00 © 2013 IEEE
DOI 10.1109/ACPR.2013.181

for improving accuracy of pedestrian tracking [5] and can be used as a cue for finding social groups [6].

Comparing with these works, our interests are in a role of group members, leadership in particular. If we can successfully estimate a leader in a group, we can get close to a deeper understanding of human activities. Our contribution in this paper is proposing a method to measure the influence of pointing actions as a cue for leadership estimation and demonstrating its potential using actual videos. Although this is a small step toward our ultimate goal, we believe this is an important milestone.

II. Measuring Influence of Pointing Action

This section describes the core of this work: how to measure the influence of pointing actions. As shown in Sec. II-A, we first segment a walking trajectory and then compute the influence measure for each pointing action. Either processes will be described in Sec. II-B and Sec. II-C respectively.

A. Basic Idea

Our proposed method is based on the intuition illustrated in FIg. 3. Suppose a leader gives an instruction for another member of a group to move on to a destination. We often observe pointing actions by the leader to indicate the goal (1 in the figure) and other members change their walking direction to follow the instruction (2 in the figure).

The proposed method measures how much influence the pointing action has by computing the consistency between the direction of the pointing and the walking direction after being changed by the pointing. To this end, we first segment a walking trajectory to find out the transition of the trajectory caused by the instruction; this process will be discussed in Sec. II-B. After the segmentation of trajectory, we compute the consistency between the walking direction and the pointing direction. as will be described in Sec. II-C.

In the following discussion, we assume that pointing actions and their directions have already been detected as shown in Fig. 4. Hereinafter, we denote a pointing action d_i is occurring at time t_i and its direction is $\mathbf{v}_p$.

B. Trajectory Segmentation

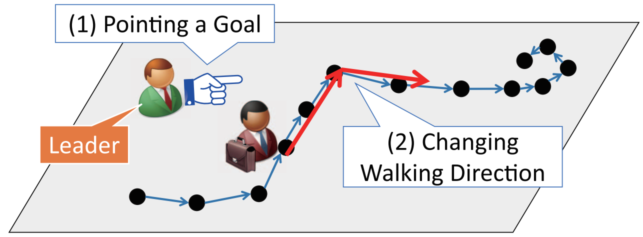


Fig. 3. Measuring Influence of Pointing Action

Figure 5 illustrates how we divide a trajectory into several segments. Black circles in the figure denote locations of a person walking into a scene.

At the first step of segmentation, we find out points in the trajectory corresponding to when he/she is stopping. We simply compute a velocity $\mathbf{v}_t$ at each time t and extract the point if $||\mathbf{v}_t||$ is smaller than a certain threshold T_v as shown in Fig. 5 (2).

These points should be removed before the next process. At the next step we will compute and use the direction of a velocity to measure the consistency between a pointing direction and a walking trajectory. Because the direction of a velocity is unstable when its magnitude is low, these points should be removed as a preprocessing.

In order to segment a trajectory, we compute two velocity values at each time t, that is, $\mathbf{v}_t^a$ and $\mathbf{v}_t^b$. The velocity values are defined as:

$$\mathbf{v}_t^a = \mathbf{p}_{t+\Delta t} - \mathbf{p}_t, \quad (1)$$
$$\mathbf{v}_t^b = \mathbf{p}_t - \mathbf{p}_{t-\Delta t}, \quad (2)$$

where let $\mathbf{p}_t$ be the position of a person in a 3D world coordinate system. Using the two velocity values, we compute how the direction of the velocities has changed by using $C_t = \mathbf{v}_t^a \cdot \mathbf{v}_t^b / ||\mathbf{v}_t^a|| || \mathbf{v}_t^b||$. When C_t is smaller than a threshold T_C, we extract the point t as a transition point of two different segments.

In Fig. 5 (3), a trajectory is divided into three segments and each segment is colored by different color. This is a very simple process but we have confirmed this yields sufficient good segmentation results for actual video scenes.

C. Influence of Pointing Action

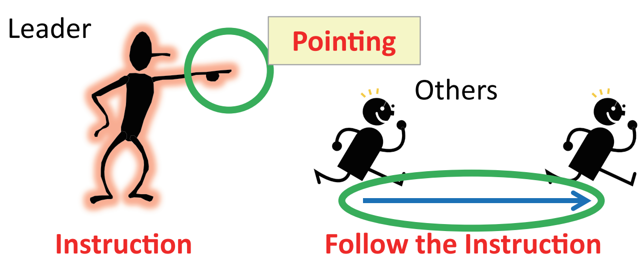


Fig. 2. Pointing Action and Leadership Estimation

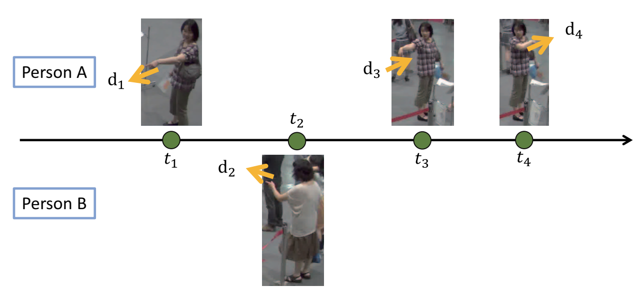


Fig. 4. Pointing Actions while Walking

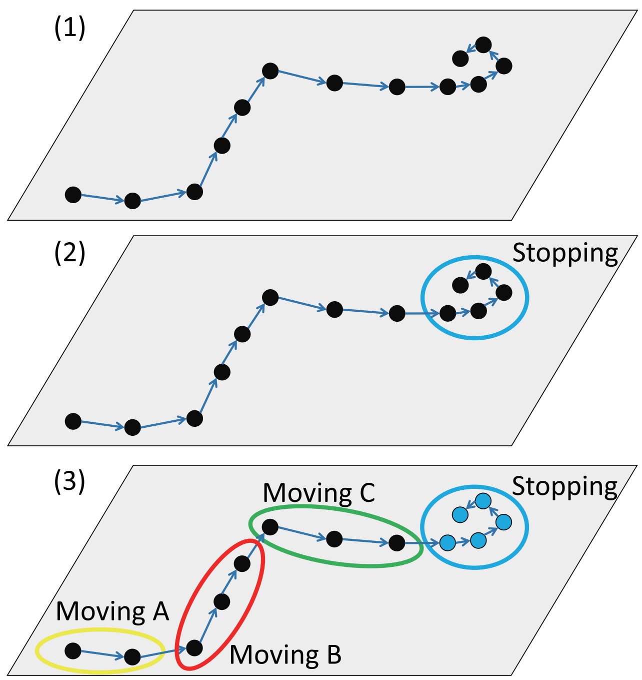


Fig. 5. Trajectory Segmentation

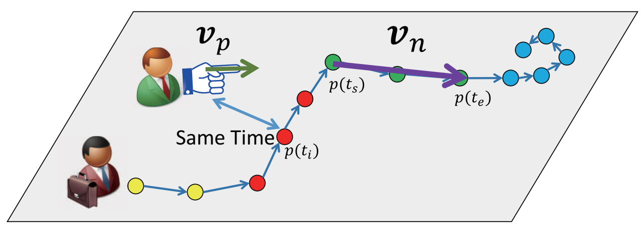


Fig. 6. Computing Influence of Pointing Action

As mentioned earlier, we assume that an influence of a pointing action can be observed how the trajectory of a person following the instruction changes. More specifically, we compute the difference in direction between a pointing action and a segment of a moving trajectory. The segment is chosen by two criteria: (1) that is temporally close to the pointing action but (2) that does not contain a point of the same time as the pointing action. A typical example is shown in Fig. 6, where we compare two vectors $\mathbf{v}_p$ and $\mathbf{v}_n$. The former denotes the direction of a pointing action and the latter does the walking direction after the transition caused by the instruction.

In order to compare the two vectors, we simply compute the dot product of them. And we sum up them among people who receive the instruction as:

$$E_K = \frac{1}{m}\sum_{n=1}^{N}(\mathbf{v}_p \cdot \mathbf{v}_n), \tag{3}$$

where let N be the number of people who receive the instruction. Hereinafter E_K is called as an IPA (Influence of Pointing Action).

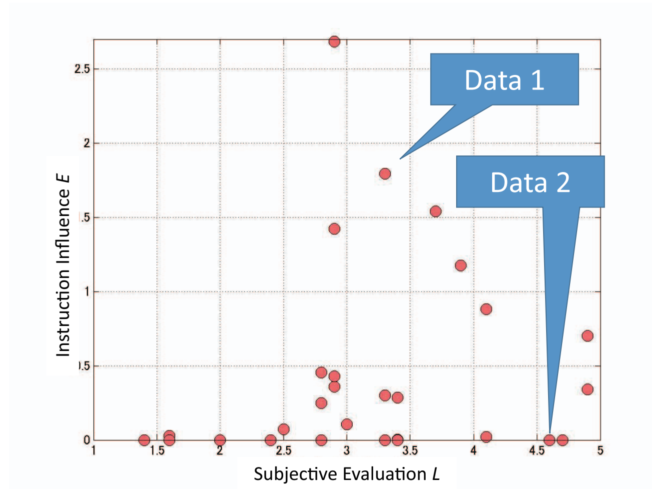


Fig. 7. Relationship Between Instruction Influence E and Subjective Evaluation L

A person is likely to do pointing actions several times while walking. We compute the maximum of IPAs calculated for each pointing action as:

$$E = \max(E_k). \tag{4}$$

Through observation we find that there are several kinds of pointing actions: pointing as an instruction is just one of them. If we sum up IPA values over all pointing actions, it contains small values even if a person is truly a leader because he/she sometimes do pointing just for indicating a certain point and IPAs for such pointing actions become small. To avoid this we take the maximum of IPAs to represent how large influence a person has. We call E *Total Influence of Pointing Action* (TIPA for short).

III. Experiment

In order to see how the proposed TIPA measures the influence adequately and it can be used for estimating leadership in a group, we conduct an experiment using an actual video captured in a science museum. In the experiments, we compute TIPA E defined in Equation 4 for each person appeared in a scene. We also conduct a subjective evaluation of a leadership for each person. Its details will be discussed in Sec. III-A.

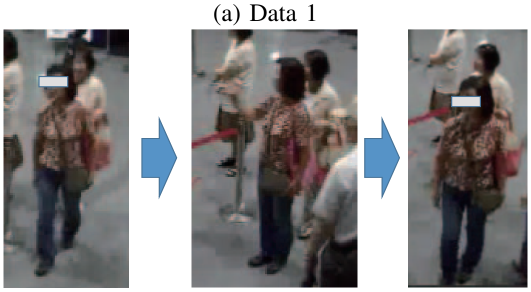
Here we should note that all moving trajectories are given by hand: all trajectories are sufficiently accurate. Pointing actions and their directions are also given by hand as mentioned in the earlier. We picked up 18 groups containing 45 people in total. 45 pointing actions by the people are observed.

A. *Subjective Evaluation of Leadership in Video*

To see whether there is a relationship between computed TIPAs and actual *leadership* of each person, we asked 10 persons to check the video and to grade the likelihood of a leader using a scalar value L from 1 to 5, in which 5 means that a person in most likely to be a leader.

Fig. 7 shows the relationship between TIPA E and the subjective score L. The vertical and horizontal axes denote E

(a) Data 1



(a) Data 2

Fig. 8. Typical Examples

and L respectively. We can see from this graph, when L takes small values, E also becomes small. On the other hand, when L is large, E can be both small and large values. This will be investigated in the next section.

B. Comparison between Computed Influence and Subjective Leadership Score

Fig. III-B(a) corresponds to ”Data 1” in the graph. A mother indicated an exhibition for a daughter to view it and the daughter followed the suggestion and moved towards it. This is one of the most typical actions which match our assumption. It is quite natural that both E and L become large.

On the other hand, Fig. III-B(b), corresponding to ”Data 2” in the graph, is an example of cases which do not match our assumption. A woman leading a group, which consist of two women, indicated a display item and talked with the other woman but they continue moving in another direction. This pointing action just indicates an object and does not mean an instruction of a leader. In such cases, it is natural to get small TIPA values.

As mentioned above, there are several types of pointing actions. Our proposed method successfully measure the influence of pointing actions if the actions are done for instructing other people. If the pointing actions have other meaning like just indicating something, our method cannot work well. This is a quite natural consequence.

IV. Summary

This paper proposes the method which measures the influence of pointing actions for estimating leadership among group members. The method uses moving trajectories and the timings and directions of pointing. We define the Influence of Pointing Action (IPA) and Total IPA (TIPA) to measure the influence.

Comparison between the computed TIPAs and subjective leadership scores demonstrates that our measures has successfully estimated the influence and has a correlation with the subjective scores when people’s actions match our assumption. Our future work is to make it possible to treat various kinds of pointing actions and to extract useful information from them.

References

[1] W. Choi, K. Shahid, and S. Savarese, “What are they doing? : Collective activity classification using spatio-temporal relationship among people,” in *9th International Workshop on Visual Surveillance (VSWS09)*, 2009, pp. 1282 – 1289.

[2] T. Kaneko, M. Shimosaka, S. Odashima, R. Fukui, and T. Sato, “Consistent Collective Activity Recognition with Fully Connected CRFs,” in *ICPR 2012*, 2012, pp. 2792–2795.

[3] W. Choi and S. Savarese, “Learning Context for Collective Activity Recognition,” in *CVPR 2011*. 2011, pp. 3273–3280.

[4] M. R. Amer and S. Todorovic, “A chains model for localizing participants of group activities in videos,” in *ICCV 2011*, 2011, pp. 786–793.

[5] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, “You’ll never walk alone: Modeling social behavior for multi-target tracking,” in *ICCV 2009*, Sep. 2009, pp. 261–268.

[6] M. Cristani, L. Bazzani, G. Paggetti, A. Fossati, D. Tosato, A. D. Bue, G. Menegaz, V. Murino, Andrea Fossati, and A. Del˜Bue, “Social interaction discovery by statistical analysis of F-formations,” in *British Machine Vision Conference (BMVC)*, 2011, pp. 23.1–23.12.