報告番号	※甲　　　第　　　　号

主　論　文　の　要　旨

論文題目　Human Subject Segmentation in Images with Complex

Background（複雑背景における人物セグメンテーション）

氏　　　名　POURJAM Esmaeil

論　文　内　容　の　要　旨

Human body segmentation has many applications in a wide variety of image processing tasks from intelligent vehicles to entertainment. A substantial amount of research has been done in the field of segmentation and it is still one of the active research areas, resulting in introduction of many innovative methods in literature. Still, until today a method that can automatically segment human subjects in different kinds of situations and with good accuracy, has not been introduced yet.

For a useful segmentation system to be realized, the following several problems that are thought to have more importance and effect 1) Shape variations due to human body movements, 2)Shape variations due to human wearing different clothes, 3) Variation in color and texture of the clothing, 4) Complexity of the scene.

The articulation problem can be considered as the most important problem. Because, the human body is made of multiple links and joints which leads to many variations in the general shape of the body during different movements in various kinds of situations. It is said that for completely modeling the human body a model with at least 20 degrees of freedom (20DOF) is necessary. This leads to a very sophisticated model which even the creation would prove to be very difficult. As a result, different approaches are usually selected instead of explicitly modeling the body ranging from 3D generative multi joint models to simple multi joint stick-man models. Also ,combination between the movements and change in the shape caused by human subjects wearing different types of clothes which can affect their general shape (rain coats, coats, T-shirts, normal shirts, and so on) adds to the complexity of the model.

Aside from this it can be said that the humans are the only species that can have numerous changes in their outer appearance by wearing different kinds of clothing. There numerous choices for humans in the clothes they use and even a single human changes his/her clothes in different situation and times. As a result a vast combination of colors and texture is created which again makes the task of modeling more problematic and sophisticated.

Complexity of the scene that a human subject is being recorded in is also one of important problems for a practical segmentation system. Even the simplest real-world cases for human eyes has proved to be quite challenging for the computer to understand and differentiate. This means that finding a model for the background will also become difficult. Especially if we use a moving camera the problem will ascend to another level of difficulty since now we need a dynamic model to cope with the situation.

Many of the recent methods try to use the graph-cut framework to solve the segmentation problem. Although powerful, these methods usually rely on a distance penalty term (intensity difference or RGB color distance). This term does not always lead to a good separation between two regions. For example, if two regions are close in color, even if they belong to two different objects, they will be grouped together, which is undesirable. Also, if one object has different parts with different colors, e.g. humans wearing clothes which have different colors and patterns, different parts will be segmented separately. Although this can be overcome by multiple inputs from user, the inherent problem would not be solved.

This thesis will try to address the above mentioned problems to propose automatic segmentation algorithms that can cope with these situations. For overcoming the changes in the shape and color we propose a method using the conventional Statistical Shape Models (SSMs) and Grab-cut segmentation algorithm. SSM uses the statistical framework to model changes in the shape of an object in Eigen space and by analyzing it and finding the most dominant changes it can even try to generate some new shapes with the properties of the original one, while the Grab-cut segmentation algorithm tries to learn image color distributions and segment the foreground object based on that in an Markov random fields frame work. By connecting these two methods through a feedback system it becomes possible to propose a coarse to fine scheme for model generation and refinement which in combination with Grab-cut leads to accurate segmentation results for human subjects.

To cast aside the color/texture problem we propose a system which makes use of a human probability map, super-pixels and Grab-cut framework. The main idea comes from jigsaw puzzle game. If we divide the image into regions based on their color/texture, each part of the human body then, becomes like a piece of puzzle. So we can think of an image as a puzzle with multiple pieces in which the human body occupies some of them. By selecting the right pieces, we can have a somewhat rough (or even fine) shape of the body and by using the Grab-cut, we can segment the human subject accurately. Following this idea, we show that not only the system becomes

automatic but also the accuracy of the system is improved. We also show that, just by using the information of a single image, it is possible to achieve segmentation results with accuracy comparable to the state-of-the-art and much better than traditional methods while having relatively simpler model. It is also good to note that this method can be used in both automatic and interactive segmentation manner.