論文内容の要旨

論文題目

An Ego-Motion Detection System Employing Directional-Edge-Based Motion Field Representations

（方向性エッジ情報を用いた動きフィールド特徴表現に基づく観察者動作認識システム）

氏名　郝　佳　（はお　じゃ）

Benefiting from the remarkable development of VLSI technology according to Moore's law, computer owns the powerful computing ability exceeding that of humans. It has taken most of the charge of computing works instead of us, and has contributed much to scientific computation. However, in spite of this success in logical and dedicated computation, computers are still not good at flexible intelligent processing, such as “recognition”. This sets the research of human-like intelligence irreplaceable tendency for the next generation of computer system. Since for humans, visual information, and especially visual motion information, is one of the most critical sources for recognition tasks, we explore machine intelligence by taking visual motion analysis as a breakthrough.

As the prerequisite of any motion analysis system, ego-motion detection that analyzes the relative motion of an observer with respect to the environment has been drawing a lot of interest in research. It plays an essential role in navigational tasks, such as automotive vehicle guidance, real-time robot control, etc. The performances of these systems are essentially determined by two major characteristics. First is the system's flexibility in adapting to disturbances, viz. illumination change, irregular observer

motion due to a variation in speed or bumping and shaking of the observer and so forth. The other is the system's operating speed that ensures a real-time response capability. Since an increase in a system's flexibility is offset by an increase in computational cost, how to compromise between these two aspects is of prime importance in building such systems.

Many contributions have been achieved for ego-motion detection by means of feature tracking, environment modeling, and so forth. Unfortunately, most of them are developed for specific missions, assuming comparative ideal circumstances such as constant illumination condition, smooth motion, etc. They therefore suffer from degradation of accuracy under severe circumstances. Moreover, most of these algorithms involve operations in frequency and spatial domain, i.e., estimating the motion models by solving some complex equations with floating-point calculations. Since the circuitry needed for the calculation of floating-point numbers is very complicated, it is difficult to implement these algorithms as a VLSI system, therefore by no means real-time practical.

On the other hand, biological systems are robust against these problems. Since the discovery by Hubel and Wiesel in the study of visual cortex of animals, it is well known that edge information plays an essential role in early visual processing. The illumination invariance of edge information makes the perception systems work well even under serious illumination conditions. In addition, for the motion interpretation stage, ``an associative architecture'' directly inspired by human physiology can more easily execute recognition process instead of the large number of numerical computations required by conventional methods, which has been successfully demonstrated by applying it to still image recognition.

The purpose of this research is to accomplish more complex recognition task, ego-motion detection, based on these biological principles, where edge information is utilized as the very basis of the system, and the ``recognition'' process of motion analysis is executed in the ``associative'' manner. The system is composed of two stages: extracting motion information from a scene, and interpreting the accumulated information. In the first stage, motion field generation, directional edge maps generated from original gray-scale images are utilized as the input. Using edge information renders the system robust against dynamic illumination variation and weak texture, while the histogram matching scheme based on edge maps for local motion detection

drastically reduces the computational cost than conventional block matching. In the second stage, motion characteristics are extracted from two perspectives of a motion field and are represented by two kinds of feature vectors. They are jointly utilized in the hierarchical classification scheme for a concise and efficient estimation of motion pattern. Multi-clue template matching in this scheme makes the algorithm robust against the motion ambiguity problem, such as distinguishing tracking and panning motions of the same direction. Moreover, by introducing a new scheme in hierarchical classification, motion field distortion due to camera shaking during video capture has also been resolved.

All operations in both the motion field generation stage and the hierarchical motion pattern classification stage are executed with integer or 1-bit calculation, namely fixed-point operation, thus it is easy to implement the algorithm with compact hardware circuitry to meet real-time need.

The ego-motion detection system has also been applied to motion estimation of hand-held devices, such as mobile phones. Digit-writing gesture recognition was taken as a target problem, where the writing stroke is recorded from a image sequence taken by a moving camera. The automatic speed adaptation capability developed in the motion detection system has enabled very robust writing stroke detection. As a result, the temporal stroke distortion due to irregular writing speed has been eliminated. Since the writing stroke is correctly reconstructed by integrating direction and magnitude of motion results, feature vector for each digit character was constructed by connecting feature distribution in each direction. As a result, handwriting gesture recognition is achieved by simple template matching. The system performance has been evaluated by digit-writing gesture recognition with irregular writing speed, different users, or cursive writing. Recognition of hand-writing Chinese characters is also attempted and the potentiality for more complicated hand-writing patterns by the algorithm has been examined. The result shows that it is possible to build higher level interfacing based on only vision information and broader use of ego motion detection to motion estimation of mobile can be expected.

In this research, an ego motion detection system employing directional-edge-based motion field representations has been developed, and successfully applied to camera motion estimation, and digit-writing gesture recognition. Usage of directional edge information renders the illumination-independent performance in motion field

generation. Hierarchical vector representation of motion field has resolved the problems of motion ambiguity and motion field distortion by simple vector processing. The introduction of the speed adaptation scheme is very effective for correct camera motion detection and therefore correct stroke generation for trajectory recognition. The algorithms proposed in this work have enabled us to build real-time response systems using the dedicated VLSI chips developed for the processing. The flexibility and the speed performance of the system simultaneously are achieved. The performance of the system has been evaluated under various circumstances and the capability of the directional edge-based motion field representation algorithm in performing robust visual motion perception has thus been verified.