報告番号	※ 甲 第 10728 号

主 論 文 の 要 旨

論文題目 Compressive Acquisition and Computational Reconstruction of Ray Space （光線空間の圧縮取得と計算的再構成）

氏　　名 要　強

論 文 内 容 の 要 旨

Ray space (light field in other literatures) has been regarded as one of the most exciting medias in image representation and computer vision. Ray space records the whole information of lights that pass through one real space, and the applications of ray space have become more popular and more attractive, which include free viewpoint image (FVI) generation for watching different perspective images, all-in-focus images generation for image refocusing, and compressive display. However, the data amount of a ray space is quite huge, and this problem is enhanced by the requirement of higher image resolution and denser sampling rate of ray space. Therefore, in this thesis, we investigate the compressive acquisition and computational reconstruction of ray space by adopting the conception of compressed sensing, which states that signal can be exactly reconstructed from incomplete measurements obtained by incoherent projection if the signal is sparse by itself or sparse in other domain. The basic target of our work is to simulate the compressive acquisition process of ray space, and evaluate the relation between the reconstruction quality and the number of sensed measurements. The simulation results can provide a good reference in the design of optical acquisition system. Based on this target, we also attempt to attain better reconstruction quality at a fixed sensing ratio by proposing new methods, which are regarded as the main contributions of this thesis.

The first contribution relates to the combination of sparse coding and reconstruction of the ray space. After establishing the acquisition model and obtaining the compressively sampled measurements of a ray space, in the reconstruction procedure, we propose to reconstruct the ray space from compressively sampled measurements by exploring sparsity of the ray space, where epipolar plane image (EPI) is adopted as the processing unit. In order to obtain sparser representation of EPI, two types of dictionaries are proposed. Normally, Gabor functions can be used to encode the motion compensated frames in video processing, and similarly EPI actually reflected the disparity information between each free viewpoint image (FVI). Therefore, the dictionary is carefully designed using Gabor function by tuning different parameters, so that the generated atoms in 2D Gabor dictionary can match the features of EPI. As a result, sparser representation of EPI can be obtained. In the experiments, we also take two other orthogonal bases to make comparison of the reconstruction quality. The results show that better reconstruction quality is achieved by 2D Gabor dictionary. The success of adopting overcomplete dictionary (whose number of atoms exceeds the dimension of the dictionary space) in representation and reconstruction of ray space promotes us to continuing exploring better dictionary by more sophisticated method.

Therefore, another dictionary is proposed by adopting dictionary learning method, and the dictionary can be adaptively amended from a set of training data, thus the atoms obtained from dictionary learning can be shaped to grasp features of EPI. In the experiments, we compare the 2D Gabor dictionary, the learned dictionary, and another orthogonal basis. The best reconstruction quality of ray space is achieved by the learned dictionary, followed by the 2D Gabor dictionary, and the orthogonal basis gets the worst result.

However, the two dictionaries mentioned above sacrifice reconstruction speed to gain high reconstruction quality, and the slow reconstruction speed becomes an obstacle for real time applications. Therefore, in order to get a faster reconstruction speed while preserving the good reconstruction quality, we propose a statistically weighted model and integrate the model into optimization for sparse solution, which is regarded as the second contribution.

The second contribution attempts to handle the trade-off between reconstruction quality and speed. We find that the amplitude structure of coefficients of ray space in DCT domain also provides another piece of priori information in the reconstruction, and we propose designing a weighted matrix to reflect the structure and to integrate the structure in the reconstruction process, which is operated by L1 norm optimization. In addition, we provide a solution for the new optimization problem so that the previous optimization solver can be reused. The experimental results show that the weighted model-based method achieves better reconstruction quality and faster reconstruction speed than the conventional method, plain L1 norm optimization. Furthermore, the weighed model-based method and dictionary learning-based method achieve similar performance (much better than the conventional method) in the aspect of reconstruction quality. However, the weighed model-based method has great advantage in reconstruction speed. Therefore, the second contribution obtains faster reconstruction speed while preserving high reconstruction quality.

Based on the contributions mentioned above, the compressively sampled ray space can be well reconstructed by computational methods. Even at quite low sensing ratio, such as 10

% of the total pixels in one ray space, the main part of ray space can still be recovered. As the sensing ratio increases to around 50 %, the reconstruction quality of ray space is quite promising. In addition, considering the reconstruction time, the second contribution can greatly accelerate the reconstruction speed, because the addition of weight promotes the convergence in the iteration for estimation of sparse solution. Therefore, the huge data problem of ray space in acquisition stage is resolved, and the ray space can also re constructed well by the proposed computational reconstruction methods.