

# 多视点视差估计中基于图像块特征矢量和相似尺度的参考视点选择方法

郁梅<sup>1)2)</sup> 董海涛<sup>1)</sup> 蒋刚毅<sup>1)2)</sup> 王让定<sup>1)</sup>

<sup>1)</sup> (宁波大学信息科学与工程学院, 宁波 315211) <sup>2)</sup> (北京大学视听信息处理国家重点实验室, 北京 100871)

**摘要** 在多视点视频处理中,采用多参考视点的视差估计方法能够在很大程度上消除各个视点间的冗余性,达到对多视点视频数据有效编码压缩的目的。研究了多视点信号处理中视差估计的多参考视点选择问题,采用快速Walsh变换谱构造图像块特征矢量,并提出了基于图像块特征和相似尺度的多参考视点选择方法。实验结果证明,本文方法能在保证编码质量的前提下,减少视差估计的计算复杂度。

**关键词** 多视点视频 参考视点 图像块特征矢量 相似尺度 视差估计

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2006)12-1834-05

## Reference View Selection Based on Characteristics of Image Block and Similarity Measurement in Multi-view Disparity Estimation

YU Mei<sup>1)2)</sup>, DONG Hai-tao<sup>1)</sup>, JIANG Gang-yi<sup>1)2)</sup>, WANG Rang-ding<sup>1)</sup>

<sup>1)</sup> (Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211)

<sup>2)</sup> (National Key Laboratory of Machine Perception, Peking University, Beijing 100871)

**Abstract** Disparity estimation is used to eliminate redundancy among different views in multi-view video coding in order to obtain high compression efficiency. In this paper, characteristics of image block are extracted by using fast Walsh transformation, and a new method for reference view selection in multi-view disparity estimation is presented, based on the characteristic vector of image block and its similarity measurement. Experimental results show that the proposed method can reduce the computational complexity in multi-view video coding with almost the same quality of reconstructed images.

**Keywords** multiview video, reference view, characteristic vector of image block, similarity determination, disparity estimation

### 1 引言

多视点视频信号是由相机阵列拍摄得到的视频信号,它能提供拍摄场景中不同角度的视频信息,并可根据使用者的喜好自由选择和操作视听对象。利用多视点视频中一个或多个视点的信息可以合成任意视点的信息,达到自由切换任意视点的目的<sup>[1]</sup>。多视点视频数据随着视点数目的增加而成倍增加,

如何合理地存储和传输多视点视频中庞大的数据是目前的一个研究热点,MPEG组织也已经在3维音视频(3DAV)中提出了基于多视点视频处理的相关标准的探索研究<sup>[2,3]</sup>,以期将多视点视频推向实用化。

在多视点视频信号中各视点间存在着很大的冗余,和传统的视频处理中用运动估计来消除时间冗余类似,用视差估计的方法可以消除各个视点间的冗余,达到对多视点视频数据更好地进行编码压缩

基金项目:国家自然科学基金项目(60472100)浙江省自然科学基金青年人才基金项目(RC01057)浙江省科技攻关项目(2004C31105);浙江省自然科学基金项目(Y105577)宁波市科技局项目(2004A610001 2004A630002)

收稿日期:2005-12-13 改回日期:2006-03-03

第一作者简介:郁梅(1968~)女,教授。2000年获韩国Ajou大学工学博士学位。主要研究方向为多媒体信号处理、压缩与通信等。

E-mail: yunei@nbu.edu.cn

的目的。因而视差估计是多视点视频编码中的关键技术之一<sup>[4]</sup>,视差估计性能的优劣直接影响到整个系统的运行效率和整个多视点视频的重构质量。就多视点视频的视差估计, Kimata 等人提出了多参考视点的方法来<sup>[5-6]</sup>提高视差估计的精度,但是该方法导致计算复杂度成倍上升,大大增加了编码器的负担。Wang 等人对多视点视频处理采用了在自然场景中提取视频对象的方法<sup>[7]</sup>,虽然同时解决了多视点视频压缩和虚拟视点建立的问题,但视频对象提取技术本身还存在一定问题。Chien 等人提出的混合二步的视差估计方法在一定程度上减少了计算量<sup>[8]</sup>,Wang 还提出了基于最小冗余预测的快速视差估计方法<sup>[9]</sup>,但是这两种算法应用于多视点视频系统其编码效率还是无法满足要求。从降低视差估计的计算复杂度出发,本文提出了一种基于图像块特征矢量和相似尺度的多参考视点快速选择算法。在保证视差估计精度和编码效率的前提下,本文方法降低了计算复杂度。最后实验结果证明该方法在多视点视频处理中是有效的。

## 2 基于图像块特征和相似尺度的参考视点选择方法

图 1 给出了“Cup”多视点图像测试集分别采用 1、2、3、4 个参考视点情况下进行视差估计补偿预测的结果。用多个参考视点进行视差估计可以显著提高系统的压缩效率,但随着参考视点数的增加,计算复杂度也会随之成倍上升,而当参考视点数量增加

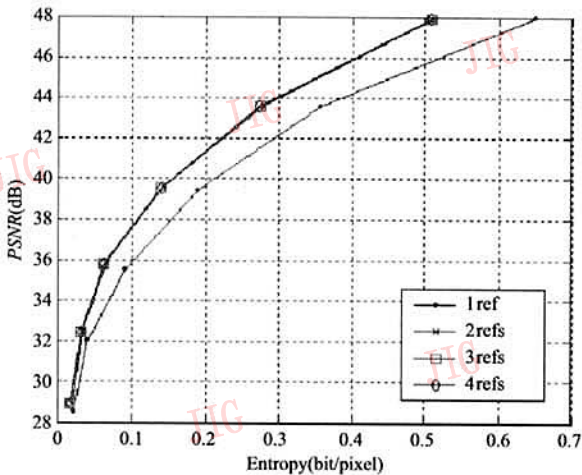


图 1 不同数量参考视点对多视点编码效率的影响

Fig. 1 Effect of the number of reference views to coding efficiency

到一定程度时其对系统编码效率的提高没有明显作用。在对“Cup”测试序列进行视差估计时,用 2 个参考视点得到的信息已满足系统编码的要求,再增加参考视点已无助于提高编码效率。因而在实际应用中选择合适的参考视点就可以很好地对多视点视频进行编码。在多视点视频处理中,采用图 2 所示的多参考视点方法能够在很大程度上消除各视点间的冗余,大大提高系统的编码性能。但系统的计算复杂度也相应成倍提高,这对多视点视频的应用而言是极为不利的。本文从降低多视点视差估计计算复杂度出发,提出了一种基于 Walsh 变换的图像块特征与相似尺度的参考视点快速选择方案,并与其他方案进行了比较。

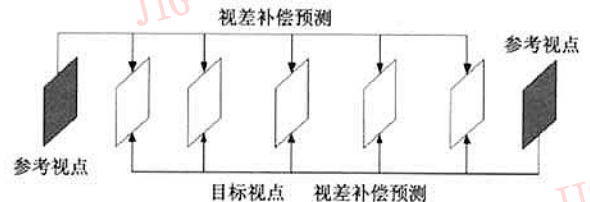


图 2 采用多参考视点的视差估计补偿预测

Fig. 2 Disparity estimation with multi-reference views

### 2.1 基于 Walsh 变换的图像块特征与相似尺度

对于 2 维离散信号  $f(x, y)$ , 其 Walsh 变换  $W(u, v)$  可表示为

$$W(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \times \prod_{i=0}^{n-1} (-1)^{[b(x) \cdot b_{n-1-i}(u) + b(y) \cdot b_{n-1-i}(v)]} \quad (1)$$

对一个  $8 \times 8$  的图像块采用式 (1) 进行 Walsh 变换可以得到一个能量相对集中的  $8 \times 8$  矩阵, 图 3 是一幅图像  $8 \times 8$  的图像块的 Walsh 变换系数统计平

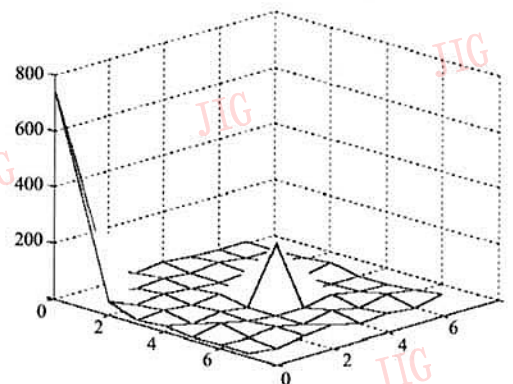


图 3 一幅图像  $8 \times 8$  图像块 Walsh 变换系数统计平均

Fig. 3 Statistic of Walsh coefficients in  $8 \times 8$  blocks of an image

均其 Walsh 系数的能量主要分布在几个特定位置。根据图 3 的统计平均分析,本文利用  $8 \times 8$  块 Walsh 矩阵  $W(u, \nu)$  中的 4 个 Walsh 变换谱系数构造一个特征矢量以表征该块的性质,即

$$T = (W(0, \rho), W(1, \rho), W(0, \mu), W(4, \mu)) \quad (2)$$

采用 Walsh 变换谱系数来构造块特征矢量是因为 Walsh 变换为正交变换,其谱系数的含义有类似于离散余弦变换相似的意义,但其计算量明显小于离散余弦变换。

根据视点中图像块特征矢量的相似距离作为相似性检测尺度,定义第  $n+p$  个块特征矢量  $T_{n+p}$  和第  $n$  个块特征矢量  $T_n$  的相似尺度为

$$D(T_{n+p}, T_n) = 0.5 |W_{n+p}(0, \rho) - W_n(0, \rho)| + |W_{n+p}(0, \mu) - W_n(0, \mu)| + |W_{n+p}(1, \rho) - W_n(1, \rho)| + |W_{n+p}(4, \mu) - W_n(4, \mu)| \quad (3)$$

根据式(3)在计算块间相似度时,不必对整个块进行 Walsh 变换,只要求计算出特定位置的 Walsh 系数即可。对于  $16 \times 16$  的宏块,采用其 4 个  $8 \times 8$  子块特征矢量平均得到。

### 2.2 参考视点快速选择算法

参考视点选择方案利用了相邻块视差信息,这里定义相邻块如图 4 所示,当前块  $c$  的 3 个相邻块  $a, b, d$  都已经完成视差估计得到相应的视差矢量和块特征矢量。

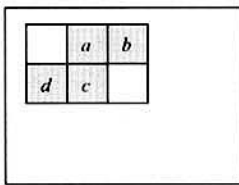


图 4 相邻块定义

Fig. 4 Definition of neighboring blocks

首先分别计算当前块  $c$  和 3 个相邻块  $a, b, d$  的相似尺度,得到一个和当前块  $c$  最接近的块(这里假设为块  $a$ )。则块  $a$  的参考视点(这里假设为左参考视点  $R_L$ )也是当前块的参考视点之一,块  $a$  的视差矢量作为当前块  $c$  在左参考视点  $R_L$  的预测视差矢量。块  $c$  在右参考视点  $R_R$  的预测视差矢量可以由左、右参考视点和当前视点的位置关系以及块  $c$  在左参考视点  $R_L$  的预测视差矢量插值得到,如图 5 所示。

用式(3)分别计算当前块  $c$  和  $c$  在两个参考视点中的预测块  $c_L, c_R$  之间的相似度距离,当块  $c$  和某

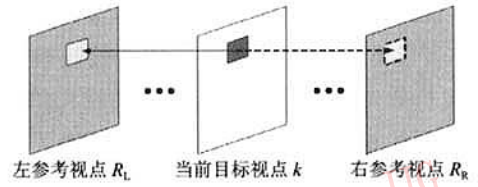


图 5 参考视点中的预测块

Fig. 5 The predicted block in reference view

个参考视点中的预测块接近到一定程度时,进行多参考视点视差估计时参考的就主要是该参考视点,而其他参考视点对视差估计的作用不大。

$$I_c = \begin{cases} \text{Ture} & D_{c,L} < D_{c,T} \\ \text{False} & D_{c,L} \geq D_{c,T} \end{cases} \quad (4)$$

$$D_{c,T} = \frac{1}{3}(D_a + D_b + D_d) \quad (5)$$

式(4)中  $I_c$  表示目标视点中块  $c$  采用的视差估计方式,  $D_{c,L}$  表示块  $c$  和参考视点  $R_L$  中预测块  $c_L$  的相似度,  $D_{c,T}$  为相似度判断阈值,表示为当前块  $c$  的 3 个相邻块与其各自的参考视点中的匹配块的相似度均值。当  $I_c$  为 Ture 时,采用参考视点  $R_L$  作为单一参考视点进行视差估计,否则采用多参考视点进行视差估计。这样,在保证视差估计精确度的前提下,减少了计算复杂度,提高了多视点编码系统的实用性。

根据上面的分析,提出基于块特征矢量与相似尺度的参考视点选择方法,步骤如下:

- (1) 计算当前块的 Walsh 变换特征矢量。
- (2) 利用得到的 Walsh 变换特征矢量分别计算和 3 个相邻块的相似尺度。
- (3) 用和当前块最相似的那个相邻块的视差矢量作为当前块在该相邻块参考视点  $R_{pred}$  的预测视差矢量  $d_{pred}$ 。
- (4) 利用左、右参考视点和当前视点的位置关系以及当前块在预测参考视点  $R_{pred}$  的预测视差矢量  $d_{pred}$  插值得到当前块在另一参考视点的预测视差矢量。
- (5) 分别计算当前块与其在两个参考视点的预测块的相似尺度。

(6) 判断和当前块最相似的预测块的相似度是否小于相似度判断阈值,如满足则只进行单参考视点视差估计,否则根据两个预测矢量在两个参考视点进行多参考视点视差估计。

(7) 完成当前块视差估计,并继续进行下一块视差估计直至处理完所有块。

### 3 实验结果与分析

为了验证本文所提出算法的有效性,对如图 6 所示的由平行相机系统获取的 Xmas、Cup、Note 等 3 个多视点图像测试集进行实验分析,每个测试集有 10 个视点,视点图像尺寸为  $640 \times 480$ ,YUV(4:2:0) 格式。其中 Xmas(由日本名古屋大学提供)和 Note 各视点间距为 30mm,Cup 各视点间距为 15mm。

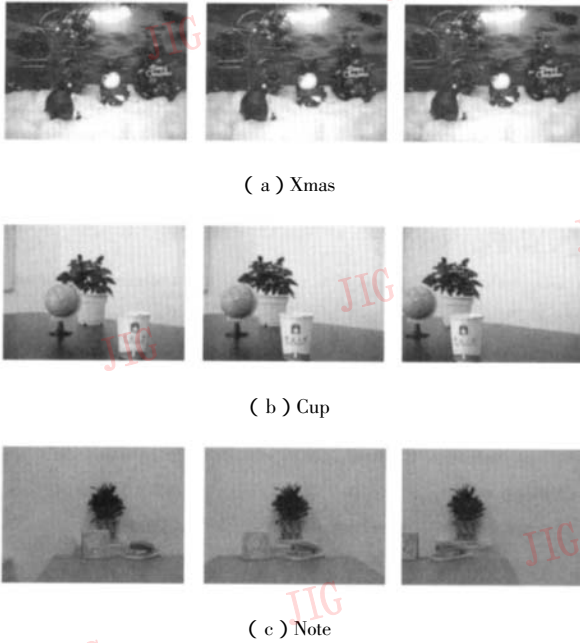


图 6 多视点测试集 10 个视点中的 3 个视点  
Fig. 6 Three views of multi-view image sets with 10 views

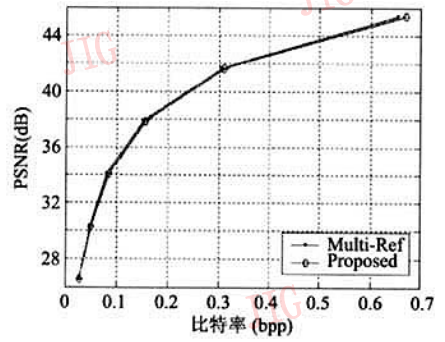
图 7 给出了 Xmas 视点图像中的 3 个块,从图中看到块 1 为边缘区域,块 2 和块 3 属于同一物体的平坦区域。显然,从图像域观察到块 1 和块 2 不相似,块 2 和块 3 是相似的。计算图 7 中 3 个不同  $8 \times 8$  块的 Walsh 特征矢量有  $T_1 = (966, 8, -477, 307)$ ,  $T_2 = (1872, 0, 0, 0)$ ,  $T_3 = (1872, 0, 0, 0)$ 。这样,



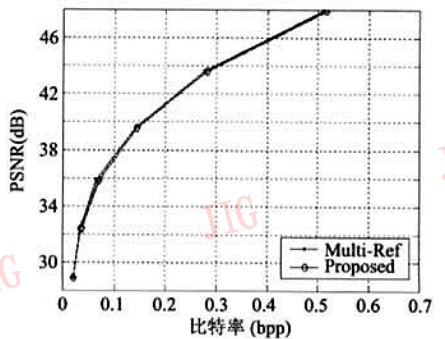
图 7 Xmas 视点图像中所选的 3 个块  
Fig. 7 Three blocks in Xmas image

块 2 与块 1、块 2 与块 3 的相似度距离分别为  $D(T_2, T_1) = 1245$  和  $D(T_2, T_3) = 0$ 。显然,用相似度距离  $D(T_{n+p}, T_n)$  计算得到块 2 与块 1 的差别较大,块 2 和块 3 的相似度距离为 0,表明这两个块完全相同。这与实际图像域的情况是完全相符的。可见用本文提出的块特征矢量相似尺度可以衡量两个块的相似程度。

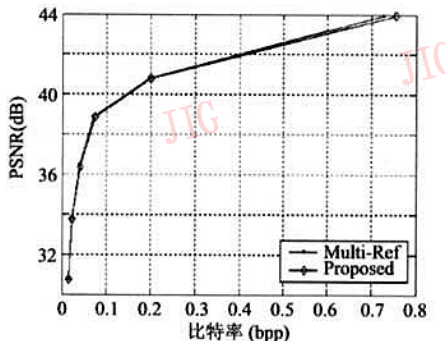
图 8 是 Xmas、Cup、Note 3 个多视点测试集进行视差估计和编码的峰值信噪比(PSNR)的比较。图中,Multi-Ref 为采用多参考视点视差估计的方法,本文只采用了最左和最右两个参考视点进行视差估计,因为 3 个以上的参考视点编码性能的提高不



(a) Xmas



(b) Cup



(c) Note

图 8 不同多视点测试集比较结果

Fig. 8 Comparison results of different multi-view image sets

明显,反而增加了计算量;Proposed 表示本文提出的基于相似尺度的参考视点自适应选择方案。实验结果表明采用本文算法,与多参考视点的编码质量基本相同。

整体的计算复杂度(即包含了视差估计、编码等全过程,而非仅仅视差估计过程本身)性能比较如表 1 所示,表中的耗时比参数表示本文所提出的参考视点自适应选择算法相对于 Multi-Ref 的两个参考视点视差估计(即未作自适应参考视点选择)时的相对消耗时间比,显然,耗时比越小越好。表中显示,在与 Multi-Ref 多参考视点的信噪比相当的情况下,本文方案通过自适应选择参考视点降低了视差估计的计算量和整个多视点图像编码过程的计算复杂度。

表 1 多参考视点视差估计和本文算法计算复杂度的比较

Tab. 1 Comparison between multi-reference disparity estimation and the proposed method on computational complexity

		算法	
		Multi-Ref	Proposed
Xmas	比特数	938 288	949 763
	PSNR( dB )	41. 67	41. 66
	耗时比( % )	100	65. 7
Cup	比特数	858 792	871 750
	PSNR( dB )	43. 59	43. 58
	耗时比( % )	100	68. 2
Note	比特数	604 341	616 588
	PSNR( dB )	40. 80	40. 78
	耗时比( % )	100	67. 1

## 4 结 论

多视点视频处理技术是 3DAV 中的核心技术,它涉及自由视点视频数据的编码、传输和存储等问

题。基于多参考视点的视差估计是多视点视频处理中的特有方法,也是其关键技术之一,能够大大消除各个视点间的空间冗余,达到对多视点视频有效编码压缩的目的。本文提出了一种基于块特征矢量与相似度判断准则的参考视点选择方法,在保证编码质量的前提下减少计算复杂度,提高编码性能。实验结果说明了算法的有效性。

## 参考文献(References)

- Nabangchang P, Fujii T. Experimental system of free viewpoint television[ A ]. In : Proceedings of SPIE , Stereoscopic Displays and Virtual Reality Systems X[ C ], California , USA , 2003 , **5006** : 554 ~ 563.
- Description of Exploration Experiments in 3DAV[ S ]. ISO/IEC JTC/SC29/WG11 N5169.
- Requirements on Multi-view Video Coding v. 2[ S ]. ISO/IEC JTC1/SC29/WG11 N6834.
- Grammalidis N , Strintzis M G. Disparity and occlusion estimation in multi-ocular systems and their coding for the communication of multiview image sequences[ J ]. IEEE Transactions on Circuits and System Video Technology , 1998 , **8**( 3 ) : 328 ~ 344.
- Kimata H , Kitahara M , Kamikura K , *et al.* Multi-view video coding using reference picture selection for free viewpoint video communication[ A ]. In : Proceedings of 2004 Picture Coding Symposium[ C ], San Francisco , USA , 2004 : 499 ~ 502.
- Kimata H , Kitahara M , Kamikura K , *et al.* System design of free viewpoint video communication [ A ]. In : 2004 International Conference on Computer and Information Technology[ C ], Wuhan , China , 2004 : 52 ~ 59.
- Wang R , Wang Y. Multiview video sequence analysis , compression , and virtual viewpoint synthesis[ J ]. IEEE Transactions on Circuits and System for Video Technology , 2000 , **10**( 3 ) : 397 ~ 410.
- Chien S , Yu S. Efficient stereo video coding system for immersive teleconference with two-stage hybrid disparity estimation algorithm [ A ]. In : The IEEE International Conference on Image Processing [ C ], Singapore , 2004 : 749 ~ 752.
- Wang Y. Fast minimal redundancy estimation for stereo image compressor[ A ]. In : Proceedings of SPIE , Electronic Imaging and Multimedia Technology IV[ C ], Beijing , China , 2005 , **5637** : 384 ~ 391.