

文章编号 1004-924X(2007)10-1622-06

优化预测运动矢量的快速运动估计算法

闫敬文^{1,2}, 余 见¹, 屈小波¹, 张晓玲¹

(1. 厦门大学 通信工程系, 福建 厦门 361005; 2. 汕头大学 智能制造技术教育部重点实验室, 广东 汕头 515063)

摘要:提出了一种优化预测运动矢量的快速运动估计算法。在对预测运动矢量研究的基础上,根据序列图像中运动矢量的中心-中值偏置分布特性和矢量间的时空相关性,结合运动矢量的相似度分析,选用中心、中值和时间相关的三个矢量作为基本预测矢量。设置相似门限来减少由三个空间相邻块预测矢量带来的大量冗余信息,对算法中关键的门限技术进行了改进。实验结果证明,本文算法对各种类型的运动序列都有很强的自适应性,在保持搜索准确度的同时,可大幅度提高运动估计的速度,其平均搜索速度是 FS 的 208 倍,明显优于 PMVFAST 的 146 倍、MVFAST 的 77 倍、DS 的 55 倍,提高了视频压缩中现有的运动估计算法的性能。

关键词:运动估计;块匹配;视频压缩;预测运动矢量;冗余信息

中图分类号:TP391 文献标识码:A

Novel fast motion estimation algorithm based on optimizing predictive motion vector

YAN Jing-wen^{1,2}, YU Jian¹, QU Xiao-bo¹, ZHANG Xiao-ling¹

(1. *Department of Communication Engineering, Xiamen University, Xiamen 361005, China;*

2. *Key Laboratory of Intelligent Manufacturing Technology, Ministry of Education, Shantou University, Shantou 515063, China)*

Abstract: A new Optimizing Predictive Motion Vector Fast Motion Estimation Algorithm (OPM-VFMEA) is proposed. Based on the center-median-biased distributed characteristic of the motion vector of real-world sequences and the high space-time correlation of adjacent block's motion vector, combined with similar analysis of the motion vectors, the center, median, forward vector are selected as the basic predictive motion vector to predict the current one, then the similar threshold is set to reduce the redundant information from the three space adjacent block motion vectors. In addition, the adaptive threshold to enable half-stop is also represented. Experiment results show that the algorithm is able to adapt to all types of video sequences and can offer a high performance of PSNR. The search speed of the algorithm is faster than that of the existing well-known algorithms. For case examined, it is about 208 times faster than that of FS in average, which is superior to that of 146 times for PMV-FAST, 77 times for MVFAST and 55 times for DS. Therefore, the algorithm improves the performance of existing motion estimation algorithms.

收稿日期:2007-04-24;修订日期:2007-07-03.

基金项目:国家自然科学基金资助项目(No. 60472081);航空科学基金资助项目(No. 05F07001)

Key words: motion estimation; block matching; video compression; predictive motion vector; redundant information

1 引言

运动估计^[1-2]是视频压缩的关键技术,广泛应用于 MPEG-1、MPEG-2、MPEG-4 和 H. 261、H. 263、H. 264 等国际视频压缩标准。全搜索法(Full Search, FS)准确度最高,但是运算量巨大,占了编码总运算量的 60%~80%,不适合于实时应用。为减小运动估计的计算量,保证运动估计准确度,在运动估计领域提出了很多快速算法^[2-15]。

早期快速算法有三步搜索法(Three Step Search, TSS)^[3]、二维对数搜索法(2-Dimension Logarithm, 2D-LOG)^[4]、正交搜索算法(Orthogonal Search Algorithm, OSA)^[5]、交叉搜索算法(Cross Search Algorithm, CSA)^[6]等,这些方法主要是通过限制搜索点的数目来减少运算量。但这些算法的初始步长较大,搜索易陷入局部最小,估计准确度不高。在此基础上,新三步搜索法(New Three Step Search, NTSS)^[7]用统计方法提出了运动矢量“中心偏移”特性。此后相继提出基于块的梯度下降法(Block-Based Gradient Descent Search, BBGDS)^[8]、菱形搜索法(Diamond Search, DS)^[9]、六边形搜索法(Hexagon-Based Search, HEXBS)^[10]等。在这些算法中,菱形搜索法由于其鲁棒性成为最常用的算法之一。但是这些算法均没有利用空间和时间相邻块运动矢量的相关性,始终采用固定的搜索模板,没有区分不同的运动类型和图像内容,估计准确度和搜索速度难以进一步提高。

考虑到序列的时空相关性,文献[11][12]提出了预测运动矢量的概念。利用预测运动矢量与当前运动矢量的较大相似性,减小了搜索区域,提高了搜索速度,减少了搜索陷入局部最小的概率,提高了搜索的准确度。文献[13]提出的 MV-FAST(Motion Vector Field Adaptive Search Technique)算法利用空间上的上、右上、左三个相邻块矢量作为当前块的预测矢量,但它仅在大运动状态下使用,没有考虑到序列的时间相关性。

PMVFAST(Predictive Motion Vector Field Adaptive Search Technique)和 APDZS(Advanced Predictive Diamond Zonal Search)^[14]方法则较充分考虑序列的时空相关性,利用时间和空间上六个相邻块运动矢量作为预测矢量,并以三个相邻块运动矢量的中值作为起始搜索的中心点,取得很好的效果,但是算法忽略了预测矢量之间的大量冗余信息。

本文结合相似度分析和运动矢量的分布特性,提出了一种基于优化预测运动矢量的快速运动估计算法(Optimizing Predictive Motion Vector Fast ME Algorithm, OPMVFMEA)。该方法选用对预测当前运动矢量最有效的三个运动矢量作为基本预测矢量,提高初始预测准确度,设置相似门限,减少由三个空间相邻块预测矢量带来的大量冗余信息,提高搜索速度。实验结果证明,本文算法搜索准确度高、搜索速度快,具有广泛的适用性。

2 预测运动矢量的冗余信息分析

视频序列具有很强的时空相关性,应用预测运动矢量能够极大地提高块匹配算法的搜索速度和准确度。早期的研究只使用空间相邻的上、右上、左边三个块运动矢量,后来进一步利用时空相关性,又增加了(0,0)运动矢量、中值运动矢量和前一帧中相同位置的矢量等。这六个预测运动矢量分别表示为 MV_T 、 MV_{TR} 、 MV_L 、 MV_0 、 MV_{median} 、 MV_P ,其中 $MV_{median} = \text{median}(MV_T, MV_{TR}, MV_L)$, MV 表示当前运动矢量。这里采用曼哈顿距离来分析矢量之间的相似性:

$$d_i = \|MV_i - MV\|_1, \quad (1)$$

其中 MV_i 为预测运动矢量, MV 为当前运动矢量, d_i 表示预测运动矢量与当前运动矢量的曼哈顿距离。定义相似度:

$$S_i = N_T / N, \quad (2)$$

其中 T 为一正整数, N_T 为 $d_i \leq T$ 的矢量数目,即与当前运动矢量的曼哈顿距离 $< T$ 的预测运动矢

量数目。 N 为总的矢量数目, S_i 表示 $d_i \leq T$ 的预测矢量 MV_i 个数占 MV_i 总数的比例, 文中将其作为预测矢量 MV_i 与当前运动矢量 MV 的相似度测度。

对一些标准运动视频序列进行统计。在统计中, 选用了多种类型的测试序列, 如表 1 所示。它们包括小幅度运动类型 CIF 格式的 Miss 序列, 中幅度运动类型 CIPR 格式的 Caltrain (512×400) 序列, 大幅度运动类型 SIF 格式的 Table tennis 序列, 以及综合各种幅度运动的编辑序列 Edit(SIF)。

表 1 预测运动矢量与当前运动矢量相似性统计结果

Tab.1 Similarity statistic results of predictive and current motion vector

d_i	Sequence	S_T	S_{TR}	S_L	S_0	S_{median}	S_P
$d_i=0$	Miss	0.60	0.60	0.67	0.66	0.66	0.79
	Caltrain	0.74	0.68	0.62	0.69	0.64	0.79
	Table tennis	0.69	0.64	0.59	0.65	0.67	0.70
	Edit	0.63	0.61	0.58	0.66	0.46	0.74
$d_i \leq 1$	Miss	0.86	0.88	0.87	0.96	0.90	0.97
	Caltrain	0.82	0.77	0.74	0.78	0.75	0.85
	Table tennis	0.75	0.71	0.68	0.72	0.74	0.79
	Edit	0.82	0.78	0.76	0.81	0.79	0.83

从表 1 中可以看到, 对于多数测试序列, $d_i=0$ 的概率一般达到 70% 以上, 即每个预测矢量约有 70% 的矢量与当前运动矢量是相同的, 因此选用上述六个预测运动矢量必然存在大量的冗余信息。通过表 1 还可以看到与当前运动矢量最相似的为 MV_P 。但是实际视频序列中存在大量的场景切换, 在这种情况下运动矢量没有时间上的延续性, 只能利用运动矢量的空间相关性。由文献 [14] 知, 运动矢量具有明显的中值 (MV_{median}) 偏移特性。同时考虑到运动矢量的中心偏移特性 [5] (MV_0), MV_P 、 MV_{median} 、 MV_0 这三个预测矢量较充分利用了运动矢量的时空特性和偏移特性, 构成了预测当前运动矢量最有效的基本预测矢量组合。

从上述分析可知, 需要消除的预测矢量冗余信息主要存在于 MV_T 、 MV_{TR} 、 MV_L 中。计算两两矢量之间的曼哈顿距离

$$d_1 = \|MV_T - MV_{TR}\|_1, \quad (3)$$

$$d_2 = \|MV_T - MV_L\|_1, \quad (4)$$

$$d_3 = \|MV_{TR} - MV_L\|_1. \quad (5)$$

取它们的最大值

$$d_{max} = \max(d_1, d_2, d_3). \quad (6)$$

由式 (6) 可知, d_{max} 越小, 表示这三个相邻块的运动矢量越相似, 由它们带来的冗余信息也就越多。由文献 [15] 知, 视频序列中约 70% 的块是小运动块。对于小运动块, 三个相邻块预测矢量必然相差不大, d_{max} 也就较小, 通过 d_{max} 设置阈值去除这些冗余信息, 可有效提高运动估计搜索速度。因 $MV_{median} = \text{median}(MV_T, MV_{TR}, MV_L)$, 当 $d_{max} <$ 某一个阈值时, 不选用 MV_T 、 MV_{TR} 、 MV_L , 否则选用。通过实验结果对比, 将此阈值选为 4。

3 基于优化预测运动矢量的快速运动估计算法

门限技术是运动估计中的关键技术。门限技术包括两个方面, 静止块的检测和搜索过程中的半路停止。本文利用已有的 SAD 值和运动矢量, 采用自适应方法去设置门限值。

设左、上、右上相邻块和前一帧对应块的 SAD 值分别为 SAD_L 、 SAD_T 、 SAD_{TR} 、 SAD_P , 静止块门限设为 T_S , 半路停止门限设为 T_H , T_S 的求取原则如下:

$$(1) \text{ 如果 } MV_T = MV_{TR} = MV_L = MV_P = 0$$

$$T_S = \max(SAD_L, SAD_T, SAD_{TR}, SAD_P). \quad (7)$$

(2) 如果 MV_T 、 MV_{TR} 、 MV_L 、 MV_P 中某一个运动矢量不为零

$$T_S = \min(SAD_L, SAD_T, SAD_{TR}, SAD_P). \quad (8)$$

为了降低噪声的影响, 将 T_S 的下限设为 256, 为了防止错误的检测, 上限设为 512。

T_H 的求取与式 (8) 相同, 区别在于上下限的不同, T_H 的上限设为 1 024, 下限设为 512。

结合 PMVFAST 算法, 本文 OPMVFMEA 算法的总体框图如图 1 所示。

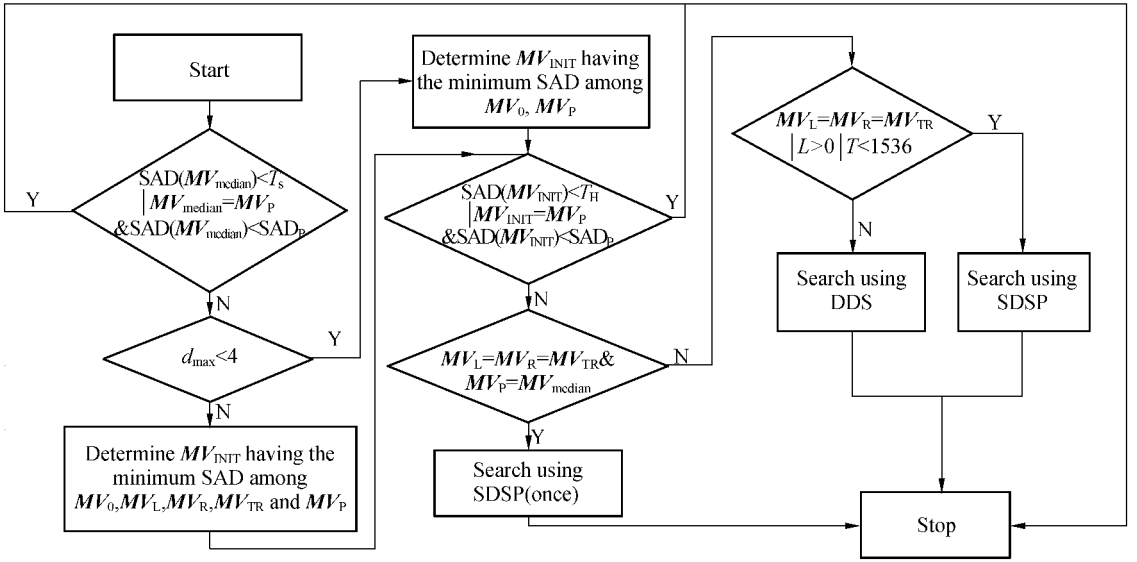


图 1 OPMVFMEA 算法的实现框图

Fig. 1 Overall structure of OPMVFMEA algorithm

算法描述如下:

(1) 使用预测 MV_{median} 。设 $SAD(MV_{median})$ 为 MV_{median} 处的 SAD 值, 如果 $SAD(MV_{median}) < T_s$ 或者 $MV_{median} = MV_P$ 且 $SAD(MV_{median}) < SAD_P$, 则停止搜索; 否则进入步骤(2)。

(2) 根据式(6)决定所使用的预测矢量。如果 $d_{max} < 4$, 则增加 MV_P, MV_0 两个预测矢量; 否则增加 $MV_T, MV_{TR}, MV_L, MV_0, MV_P$ 等五个预测矢量。找出这些预测矢量中具有最小值 SAD 的矢量, 记为 MV_{INIT} , 同时该矢量处的 SAD 值记为 $SAD(MV_{INIT})$ 。如果 $SAD(MV_{INIT}) < T_H$ 或者 $MV_{INIT} = MV_P$ 且 $SAD(MV_{INIT}) < SAD_P$, 则停止搜索; 否则进入步骤(3)。

(3) 如果 $MV_T = MV_{TR} = MV_L$ 且 $MV_{median} = MV_P$, 则仅进行一次小菱形搜索(SDSP)^[9]; 否则转到步骤(4)。

(4) 如果 $MV_T = MV_{TR} = MV_L$ 或者 $L > 0$ 或者 $T < 1536$, 认为当前块运动为小运动, 进行小菱形搜索; 否则认为当前块运动为大运动, 进行偏向菱形搜索(DDS)^[16], 直到当前最佳匹配点(SAD 最小)位于搜索模板的中心, 停止搜索。其中 $L = \|MV_{median}\|_1, T = T_H + 256, T$ 的上限设为 1 792。

4 实验结果和讨论

为验证本文算法 OPMVFMEA 的性能, 与多种经典算法如 TSS, NTSS, BBGDS, DS, HEX-BS, MVFAST, PMVFSAT 等进行了比较实验。本文中所有算法的实验仿真平台都为 Matlab 7.04 版本, 每种序列均取 150 frame, 块的大小均为 16×16 , 搜索范围为 ± 15 。实验结果如表 2 和图 2 所示, 其中 PSNR 为峰值信噪比, NSP (average Number of Search Points per block) 为每块的平均搜索点数, 分别表示搜索的准确度和速度, Speed Up 表示相对 FS 的搜索速度比。

从实验结果可看出, 本文算法具有如下特点:

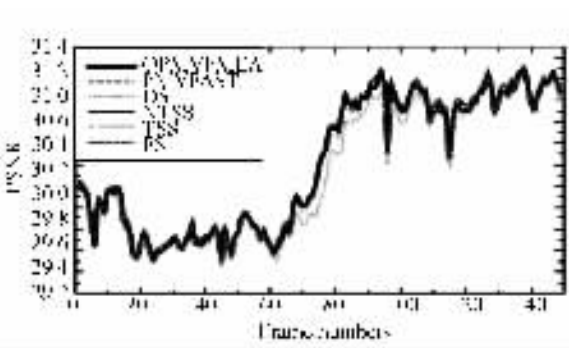
(1) 算法的搜索速度快。从表 2 中 Speed Up 值可知, 在块大小为 16×16 、搜索范围为 ± 15 的实验条件下, 平均搜索速度为 FS 的 208 倍, 而 PMVFAST 为 146 倍, MVFAST 为 77 倍, DS 为 55 倍。还可看到, NSP 都较小, 且对同一序列的任何一帧, NSP 都低于经典算法, 如图 2(b) 所示。

(2) 算法搜索的准确度高。从表 2 还可看出本文提出的方法的 PSNR 值高于绝大多数经典快速算法, 仅比 PMVFAST 低 0.01~0.03 dB。

表 2 算法仿真结果

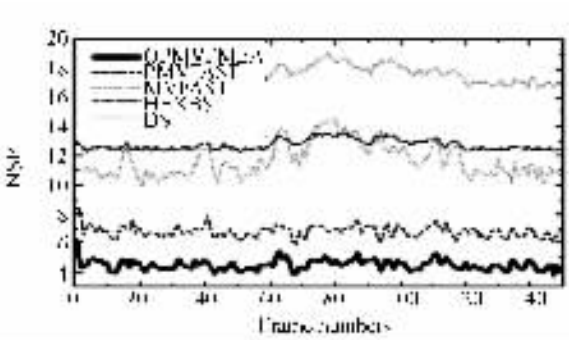
Tab. 2 Simulation results of different algorithms

Sequence	Format		FS	TSS	NTSS	BBGDS	DS	HEXBS	MVFAST	PMVF	OPMV
										AST	FMEA
Miss	CIF	PSNR	30.96	30.79	30.91	30.91	30.87	30.72	30.91	30.92	30.91
		NSP	961	41	30.88	12.15	18.03	13.01	12.27	6.53	4.49
		Speed Up	1	23.44	31.12	79.09	53.30	73.87	78.32	147.2	214.0
Caltrain	CIPR	PSNR	31.04	28.21	30.60	30.73	30.78	30.23	30.68	30.78	30.78
		NSP	961	41	23.32	13.13	16.88	13.72	13.61	6.64	4.67
		Speed Up	1	23.44	41.21	73.19	56.93	70.04	70.61	144.7	205.8
Table tennis	SIF	PSNR	28.14	26.55	27.44	27.17	27.80	27.55	27.82	27.93	27.91
		NSP	961	41	33.52	16.87	19.04	16.83	12.54	6.35	4.57
		Speed Up	1	23.44	28.67	56.97	50.47	57.10	76.63	151.3	210.3
Edit	SIF	PSNR	27.57	26.24	26.89	27.15	27.24	26.85	27.25	27.30	27.27
		NSP	961	41	27.07	12.99	16.10	12.74	10.51	6.71	4.70
		Speed Up	1	23.44	35.50	73.98	59.69	75.43	91.44	143.2	204.5



(a) 平均 PSNR 对比

(a) Average PSNR comparison



(b) NSP 对比

(b) NSP comparison

图 2 不同算法对“Miss”测试序列的搜索性能对比
Fig. 2 Performance comparison of different algorithms using “Miss” sequence

对部分序列,每帧 PSNR 值还接近于全局最优的 FS 算法,如图 2(a)所示。

(3)算法具有较强的鲁棒性。把该算法应用于各种测试序列,均能得到结论 1、2 所述结果。同样从表 2 中 Edit 序列的仿真结果也可显示算法对各种不同的变化场景都具有适用性。

值得指出的是,如果能够从理论上分析预测矢量之间的冗余,使用更好的准则去除这些冗余信息,则可以进一步提高运动估计的准确度和速度。

5 结 论

本文通过对预测运动矢量进行冗余统计分析,提出了一种基于优化预测运动矢量的快速运动估计算法。同时利用它的中心-中值偏移特性,并结合半路停止等技术,使算法的性能得到进一步提高。实验结果证明,本文算法在保持搜索准确度的同时,搜索速度明显优于其它算法,它的平均搜索速度是 FS 的 208 倍,而在相同条件下,PMVFAST 为 146 倍,MVFAST 为 77 倍,DS 为 55 倍。因此,本文算法提高了现有运动估计算法的性能。

参考文献:

- [1] 孙辉,赵红颖,熊经武,等. 基于光流模型的图像运动估计方法[J]. 光学精密工程,2002,10(5):443-447.
SUN H, ZHAO H Y, XIONG J W, *et al.*. Method of estimating image motion based on the optical flow model [J]. *Opt. Precision Eng.*, 2002, 10 (5) : 443-447. (in Chinese)
- [2] 梁燕,刘文耀. 基于起点预测的自适应交叉—准菱形运动估计算法[J]. 光学精密工程,2005,13(2):237-246.
LIANG Y, LIU W Y. Adaptive cross-quasi-diamond search algorithm based on initial search point estimation[J]. *Opt. Precision Eng.*, 2005, 13(2):237-246. (in Chinese)
- [3] KOGA T, ILINUMA K, HIRANO A, *et al.*. Motion compensated interframe coding for video conferencing[C]. *Proc. Nat. Telecommun. Conf., New Orleans, LA*, 1981.
- [4] JAIN J, JAIN A. Displacement measurement and its application in interframe image coding[J]. *IEEE Transactions on Communication*, 1981, 29(12):1799-1808.
- [5] PURI A, HANG H M, SCHILLING D L. An efficient block-matching algorithm for motion-compensated coding [C]. *Proc. IEEE CASSP-1987*, 1987: 1063-1066.
- [6] GHANBARI M. The cross-search algorithm for motion estimation[J]. *IEEE Trans. Commun.*, 1990, 38(7): 950-953.
- [7] LI R, ZENG B, LIOU M L. A new three-step search algorithm for block motion estimation[J]. *IEEE Trans. Circuits Syst. Video Technol.*, 1994, (4): 438-442.
- [8] LIU L K, FEIG E. A block-based gradient descent search algorithm for block motion estimation in video coding[J]. *IEEE Trans. Circuits Syst. Video Technol.*, 1996, (6):419-422.
- [9] ZHU S, MA K K. A new diamond search algorithm for fast block-matching motion estimation[J]. *IEEE Trans. Image Processing*, 2000, 2(9): 287-290.
- [10] ZHU C, LIN X, CHAU L P. Hexagon-Based Search Pattern for Fast Block Motion Estimation[J]. *IEEE Trans. Circuits Syst. Video Technol.*, 2002, 12(5): 349-355.
- [11] TOURAPIS A M, AU O C, LIOU M L. Fast motion estimation using circular zonal search[C]. *Proc. SPIE Symp. Visual Communication and Image Processing—VCIP'99*, 1999, (2):1496-1504.
- [12] TOURAPIS A M, AU O C, LIOU M L, *et al.*. Fast and efficient motion estimation using diamond zonal based algorithms[J]. *J. Circuits, Syst., Signal Processing*, 2001, 20(2):233-251.
- [13] HOSUR P I, MA K K. Motion vector field adaptive fast motion estimation[C]. *Int. Conf. Information, Communications and Signal Processing (ICICS 1999)*, 1999.
- [14] TOURAPIS A M, AU O C, LIOU M L. Highly efficient predictive zonal algorithms for fast block matching motion estimation[J]. *IEEE Trans. Circuits Syst. Video Technol.*, 2002, 12(10): 934-947.
- [15] CHEUNG C H, PO L M. A novel small-cross-diamond search algorithm for fast video coding and video conferencing applications[C]. *Proc. IEEE Int. Conf. Image Processing*, 2002: 681-684.
- [16] JIA H J, ZHANG L. Directional diamond search pattern for fast block motion estimation[J]. *Electronics Letters*, 2003, 39(22): 1581-1583.

作者简介:闫敬文(1964—),男,博士,教授,博士生导师,现为厦门大学信息科学与技术学院信息处理技术研究所所长,中国民主同盟会会员,中国通信学会会员,中国图像图形学会理事,主要研究方向为小波理论及其应用、图像压缩、图像处理和分析、遥感图像处理及其软件开发;E-mail:jwyan@stu.edu.cn, xdyjwen@126.com
余见(1983—),男,重庆人,厦门大学通信工程系硕士研究生,主要研究方向为图像/视频压缩。E-mail: xmu yujian@yahoo.com.cn