

基于 TMS320DM642 芯片 H.264 编码器的设计

薛 漪¹, 马伍新²

(1. 河北理工大学, 河北唐山 063000; 2. 大唐电信, 北京 100083)

摘要: 作为新一代视频压缩协议 H.264, 理论上已经证明它能比其它视频压缩协议表现出更好的性能, 更能适应无线多媒体网络多媒体的应用需求。但是由于 H.264 协议自身的复杂性以及控制的灵活性, 使得如何设计 H.264 编码器成为工程设计人员必须考虑的问题。采用具有较高运算速度的 TMS320DM642 DSP 芯片作为 H.264 编码器实现的硬件平台, 结合合适的编码器控制算法, 将有效地解决这个问题, 实现满足应用需求的 H.264 编码器。

关键词: H.264; DSP; Lagrangian

中图分类号: TP391 **文献标识码:** A **文章编号:** 1000-7024(2005)05-1158-03

Design of H.264 coder based on TMS320DM642 chip

XUE Yi¹, MA Wu-xin²

(1. Heibei University of Technology, Tangshan 063000, China; 2. Datang Telecomm, Beijing 100083, China)

Abstract: The theories have proved that H.264, a new advanced video compressed protocol, has a better performance and adaption to requirements of the application of the wireless and network multimedia better than other video compressed protocols. But because of the complexity of H.264 protocol and the flexibility of the operational control, how to design H.264 coders is the key problem that engineers and corporations that plan to implement them have to consider. If high speed DSP chips, TMS320DM642, are used as hardware platforms on which H.264 coders are based and the right algorithm of the control of the encoder is adopted, the problem should be solved well and H.264 coders that meet the requirements of the application will be implemented.

Key words: H.264; DSP; lagrangian

1 引言

H.264 协议的最新草案只给出了一个理想的解码器模型, 而把具体的编码器设计留给了工程实现者。实现者需要解决的问题有: 运动搜索的算法; 宏块的编码模式选择算法; 量化参数选择算法等等。此外, 由于 H.264 协议引入了多参考帧以及 1/2、1/4 像素精度进行运动估计, 使得 H.264 的运算复杂度也相应的提高了很多。因此, 要设计一个好的 H.264 编码器, 尤其是满足实时性要求的嵌入式 H.264 编码器, 将会面临许多困难。根据应用需求, 正确设计 H.264 编码器, 不仅可以提高其压缩性能, 而且可以减少需要的运算量。

考虑到 H.264 协议实现的复杂度, 我们决定采用 TI 公司推出的数字媒体处理器 TMS320DM642 DSP 芯片作为 H.264 编码器实现的硬件平台。本文将给出一个基于这个芯片的嵌入式 H.264 编码器的设计方案。考虑到实现的难度, 目前该编码器只实现了 H.264 协议的 Baseline 部分。

2 TMS320DM642 芯片介绍

TMS320DM642 DSP 芯片的 CPU 时钟频率可达 600MHz,

其 CPU 内部有 8 个 32 位的功能单元, 可同时执行 8 条指令。如果能充分利用这 8 个功能单元, 在完全流水的情况下, 该芯片的指令吞吐量将达到 4800MIPS。该芯片内部支持两级 Cache, 其中第一级 Cache 对开发人员来说是不可见的, 而第二级的 Cache 大小是可配置的, 芯片自动完成这两级 Cache 之间数据一致性的维护。有了这两级 Cache 的支持将使 CPU 的执行速度大大加快。同时, 64 比特宽度的 EMIF 口, 可以无缝连接到 SDRAM、SRAM 等各种存储器, 方便了存储器资源的扩展。TMS320DM642 芯片为了适应数字媒体处理的需求, 增加了一些新的功能接口: 增加了 3 个可配置的视频端口。这 3 个端口可以被独立配置, 能与各种视频 A/D、D/A 芯片进行无缝连接, 方便了硬件电路的设计。并且这些端口支持 BT.656、RAW VIDEO 等多种数字视频格式; 增加了一个 10/100 Mb/s Ethernet MAC, 并通过寄存器配置, 提供一定的网络 QoS 保证。

3 H.264 编码器硬件平台

如图 1 所示为 H.264 编码器实现的硬件平台。模拟视频信号经过 A/D 芯片 SAA7113 变换为数字信号后输入 TMS320DM642 芯片中进行处理, 其处理后的数据即编码码流可以通

收稿日期: 2004-05-09。 基金项目: 国家自然科学基金项目 (60272011)。

作者简介: 薛漪 (1959-), 女, 河北人, 讲师, 研究方向为多媒体技术和管理工程; 马伍新 (1965-), 男, 河北人, 博士, 高级工程师, 研究方向为图像处理。

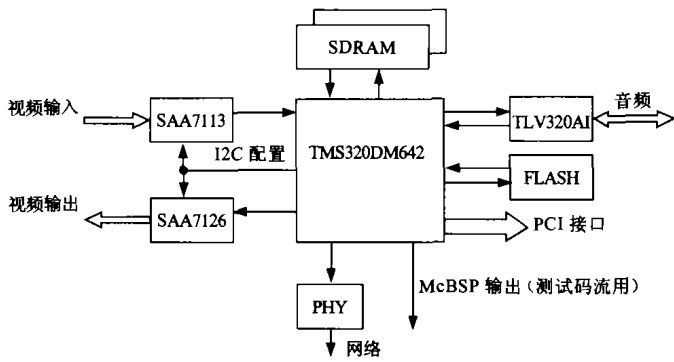


图1 H.264 编码器硬件平台

过串行通信口 McBSP 进行直接输出, 以便调试检查。或者可以通过以太网接口输出, 进行网络传输; 还可以通过 TMS320DM642 芯片的 PCI 接口输出, 把处理后的结果交给 PC 进行进一步处理。同时, 本地的重构图像可以通过 TMS320DM642 芯片输出给视频 D/A 芯片 SAA7126 后显示输出。考虑到多媒体应用的需要, 还需要对音频信号进行处理, 我们只需要将音频 A/D、D/A 芯片 TLV320AIC11 与 TMS320DM642 芯片的 McASP 口相连就可以了。

4 H.264 编码器固件设计

如图 2 所示, DSP 上 H.264 编码器固件程序需要 4 个软件模块: EDMA 数据接收处理模块、编码数据输出控制模块、本地视频显示控制模块以及核心模块 H.264 编码模块。

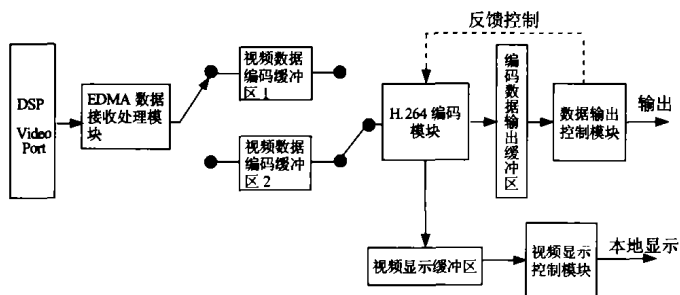


图2 H.264 编码器总体软件框图

其中 EDMA 数据接收处理模块以乒乓方式轮流从 DSP 的 VIDEO PORT 接收采集到的视频数据, 并把采集到的一帧 CIF 格式的数据按 Y 分量、Cb 分量、Cr 分量分开存放的方式放入视频数据编码缓冲区中。等待 H.264 编码模块对其进行处理; 而数据输出控制模块则负责对 H.264 编码模块输出的变速率码流进行整形, 保证以基本恒定的速率输出编码码流; 视频显示控制模块, 利用编码器中存放的参考帧, 负责本地重构图像的同步显示, 以便编码器的调试; H.264 编码模块是整个编码器的核心, 主要完成按照 H.264 协议的要求对视频数据进行实时压缩。下面将着重介绍这个模块的设计。

4.1 H.264 编码模块数据区分配

根据 Thomas Weigand 的论文, 采用 Lagrangian 技术作为该 H.264 编码器的操作控制算法。即利用 Lagrangian 公式 $D+\lambda * R$ (以下简称 R-D 公式) 进行宏块编码模式选择及运动向量的选择。其中 $\lambda = 0.85 * 2^{(QP-12)/3}$, QP 为 H.264 量化参数。

由于该 H.264 编码器是基于 TMS320DM642 DSP 芯片的,

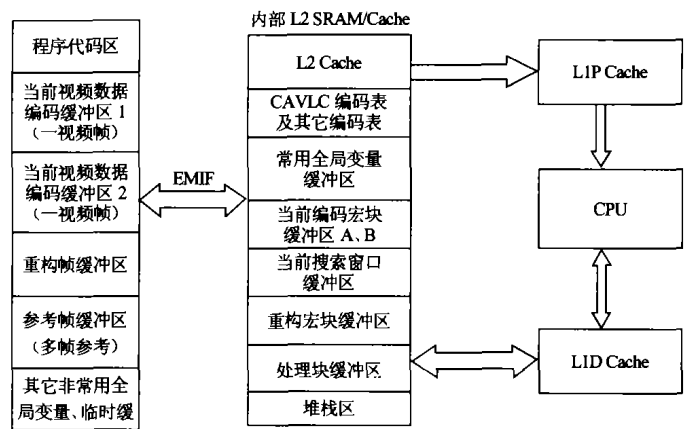


图3 TMS320DM642 内部数据区分配

为了充分发挥该芯片的性能优势, 在对 H.264 编码模块的数据区进行规划时, 必须结合该 DSP 体系结构进行考虑, 否则将会对程序优化造成困难, 降低程序执行效率。如图 3 所示, 为 H.264 编码模块在该 DSP 芯片中的数据区分配情况。对一些需要较大存储空间的数据, 如当前正在编码的帧、参考帧缓冲区等放于外部 SDRAM 中。当需要对其中的某个宏块进行编码时, 可以通过芯片的 EDMA 功能以乒乓方式轮流放入 L2 SRAM 中的当前编码宏块缓冲区 A 或 B 中, 以加快在编码该宏块时 CPU 读取该宏块数据的速度。另一方面, 在进行 INTER 编码时, 也需要通过 EDMA, 把以该宏块或块的一个预测位置为中心的一个 32×32 的搜索窗口从对应的参考帧中取出放入当前搜索窗口缓冲区中, 以便程序在该窗口内进行运动搜索。而当前的重构宏块, 先暂时存放于 L2 SRAM 内部的重构宏块缓冲区中, 以便按 Lagrangian 公式对该宏块进行 R-D 值计算。只在该宏块编码完成后才把宏块的重构数据写入外部 SDRAM 中重构帧缓冲区的对应位置。待整个帧编码完成后, 就可以对重构缓冲区中的重构帧进行 Deblock 操作, 得到经过滤波处理后的一帧图像并放于参考帧缓冲区中。

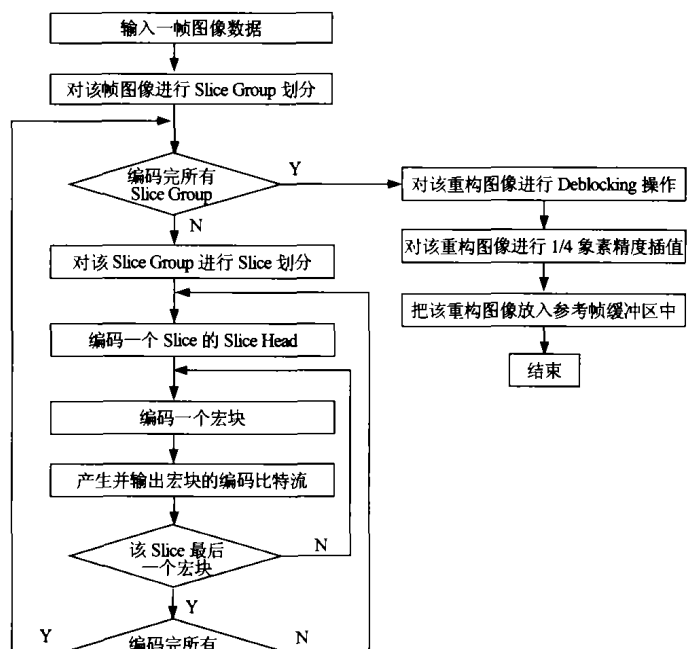


图4 H.264 编码器编码主流程

4.2 H. 264 编码器编码流程

如图 4 所示为 H.264 编码器主流程。对输入的一帧图像

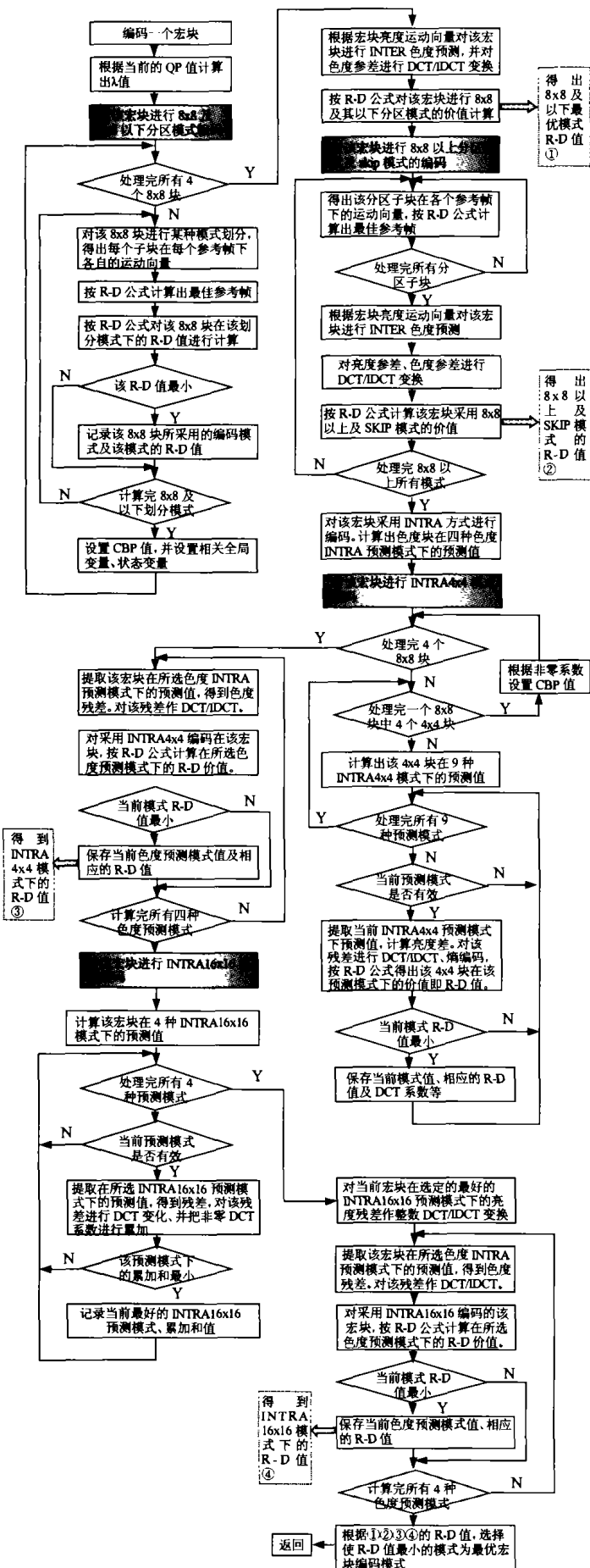


图 5 编码一个宏块流程

首先要确定该图像是编码为 I-Frame 还是 P-Frame; 其次需要对该图像进行要求的 Slice Group 划分、Slice 划分(目前只实现了一个 Slice Group 划分为一个 Slice)。经过以上步骤,每个宏块所属的 Slice 也就确定了,也就可以对每个宏块进行编码了。

如图 5 所示,对宏块进行编码是整个软件最核心也是最复杂的部分。其中图中的 DCT 指的是 H.264 的整数 DCT; R-D 值指的是按 Lagrangian 公式计算出来的值。计算每种编码模式的 R-D 值的过程大体相似,都要经过以下几个步骤:①计算亮度失真度:亮度失真度=(亮度宏块的重构值-亮度宏块的原始值)²;②计算 U、V 色度失真度:UV 色度失真度=(色度块的重构-色度块的原始值)²;③得到该宏块的总失真度:总失真度 D=亮度失真度+色度 UV 失真度;④利用熵编码函数,对该宏块的所有语法元素进行编码,包括对编码模式的编码、运动向量的编码、参考帧的编码及 DCT 系数的编码。统计总共花费的比特数 R;⑤得到 R-D 值 = D+λ*R。选择使 R-D 值最小的模式为宏块的最优编码模式。

由于 TMS320DM642 为定点 DSP 芯片,而 R-D 值的计算涉及浮点运算,这就要求我们进行相应的处理。其处理方法是采用 Q16 格式用一个 32 比特整数来近似表示 λ 值。例如当量化参数 QP=28 时,按照前面给出的公式计算得到此时的 λ=34.269 85256。我们用一个 32 比特整数来表示这个浮点数,约定小数点位于第 15 比特之后(从 0 比特开始),那么这个浮点数可以表示为 0x00224514,即 34.26983643。有了这种表示法,当对 λ*R 进行乘法运算时,就可以使用定点乘法运算指令 MPYLI,直接进行运算产生一个 48 比特的乘法结果,接下来只需把 D 的值左移 16 比特(小数点对齐),就可以直接与乘法结果相加,最终得到小数部分精度为 16 位的 R-D 值。具体在 DSP 实现时,为了避免繁琐的计算,可以先把每个 QP 值对应的一个 λ 值及其表示这个 λ 值的 32 比特整数计算出来,制成一张表。在程序运行当中时,利用查表法就可以得到对应 QP 值下需要的 32 比特整数 λ。

为了减少搜索时间,该编码器采用 3 步搜索法作为帧间预测编码时使用的运动估计算法。并结合 Lagrangian 技术对运动向量进行选择。即不仅考虑编码块与目标块的匹配程度,也要考虑编码该运动向量需要的比特数,以提高编码器的压缩性能。

5 结论

我们用几个测试视频序列对该 H.264 进行了性能测试。如表 1 所示,该 H.264 编码器表现出较好的压缩性能。与标准 H.263 编码器的性能对比情况来看,该 H.264 编码器也表现出更好的性能。但是该 H.264 编码器只实现了协议的 Baseline 部分,而协议的其他部分,由于其复杂性,则需要工程实现人员进行进一步考虑。

表 1 H.264 编码器性能测试

视频序列	格式	帧率(Hz)	亮度 PSNR (dB)	码率(kbit/s)
Akiyo	CIF	15	35.02	21.67
News	CIF	15	32.67	57.84
Clare	CIF	15	36.46	21.77
Container	CIF	15	31.38	32.48

(下转第 1163 页)

新其出现的次数。如果其出现次数已经大于前一个(LEVEL、RUN)的出现次数,则交换这两个(LEVEL、RUN)在表中顺序。

步骤 4 如果码表中的最大出现次数过大,为了防止溢出,重新执行步骤 2。

由于上述码表动态重组算法是基于过去帧的符号概率,因为解码端也可通过上述算法得到码表,动态码表不用传到解码端,因此实现此算法不用在生成码流中传送多余的信息。

2.6 其它问题

同样,在模式判决过程中,如果使用哈达玛变换,那么哈达玛变换也应该和自适应块变换类似,变换块大小要和预测块大小结合起来,这样才能更精确地反映各个块模式的优劣。

在本方案中码流的句法不需要做太大的改变。在帧间编码时,由于变换块大小和运动估值块大小一样,而运动估值的块大小已经是句法的一部分,所以不用传送变换块的大小信息。但在帧内预测时原来的两种块模式增加到现在的 7 种块模式,需要对新增的帧内预测模式进行编码,之后变换块大小就不用传送了。

3 仿真结果

整套方案包括基于简化自适应块变换的帧间编码和帧内编码、量化、游程编码和自适应通用变长编码已经在联合视频小组提供的模型版本 1.0(JM1.0)上实现,并根据文献[6]中提供的通用编码参数在一系列标准测试序列上测试了其性能,所得率失真曲线如图 4 所示。

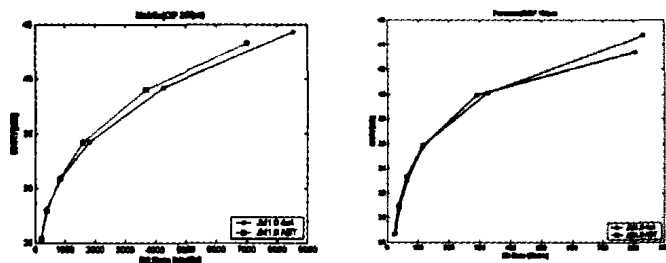


图 4 Mobile (CIF), Foreman (QCIF) 的率失真曲线

根据图 4 中的率失真曲线,我们使用了文献[7]中建议的方法估算平均信噪比提高,平均比特率节省和最高信噪比提高。当源序列为 30 帧/秒的 Mobile(CIF)时,平均信噪比提高为 0.4392dB,平均比特率节省为 8.1044KB,最高信噪比提高为 1.03dB;当源序列为 10 帧/秒的 Foreman(QCIF)时,平均信噪比提高为 0.1086dB,平均比特率节省为 0.6250KB,最高信噪比提

高为 1.99dB。与原来以 4×4 整数余弦变换为基础的 JM1.0 相比较,使用了本套算法的编码器性能提高十分显著。就平均特性来说,对分辨率较大的 CIF 序列,在整个码率范围性能都有比较均匀的提高,而对分辨率较小的 QCIF 序列,性能有波动的情况出现;而就局部特性来说,本算法在码率较高时性能提高很大,而在码率较低时性能提高不是很明显,甚至会出现些微的性能降低。

4 结论

实验和应用表明,与原来的 H.264 编码算法相比,本算法具有复杂度增加不大、在码率较高时性能提高十分明显、而在码率较低时效果不是很明显等特点。在实际应用中,可以根据源图像序列的分辨率和应用所要求的复杂度选择是否使用本算法,或者配合速率控制,在动态调整量化参数的同时自适应的选择是否使用本算法。

参考文献:

- [1] Wiegand T, Sullivan G. H.264 working draft number 2[EB/OL]. 2002-01-29. <http://ftp3.itu.int/av-arch/video-site/h26L/jwd2r2.zip>.
- [2] Hussain Y, Farvardin N. Adaptive block transform coding of speech based on LPC vector quantization[J]. IEEE Trans Signal Processing, 1991, 39(12):2611-2620.
- [3] Cham W K. Development of integer cosine transforms by the principle of dyadic symmetry[J]. IEEE Proc Communications, Speech and Vision, 1989, 136(1):276-282.
- [4] Wien M, Dahihoff A. Intra coding using variable block sizes[EB/OL]. 2001-11-26.ftp://standards.pictel.com/video-site/0112_Pat/VCEG-O31.doc.
- [5] Choi W I, Jeon B W. Dynamic UVLC codeword remapping with fixed re-association table for H.26L[A]. Picture coding symposium[C]. Seoul, Korea, 2001.167-170.
- [6] Sullivan G. Recommended simulation common conditions for H.26L coding efficiency experiments on low-resolution progressive-scan source material[EB/OL]. 2001-09-27. ftp://standards.pictel.com/video-site/0109_San/VCEG-N81.doc.
- [7] Bjontegaard G. Calculation of average PSNR differences between RD curves [EB/OL]. 2001-03-26.ftp://standards.pictel.com/video-site/0104_Aus/VCEG-M33.

(上接第 1160 页)

参考文献:

- [1] JVT-100, Study of Final Committee Draft of Joint Video Specification ITU Rec. H.264/ISO/IEC 14496-10 AVC) [S].
- [2] Ti, TMS320DM642 Video/Imaging Fixed-Point Digital Signal Processor [CP/CD].
- [3] Thomas Weigand, Bernd Girod. Parameter selection in lagrangian hybrid video coder [R]. Thessaloniki, Greece: ICIP2001, 2001.
- [4] Ohta Y, Kanade T. Stereo by intra- and inter-scanline search using dynamic programming[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1985, 7(2):139-154.
- [5] Friedman G L. The trustworthy digital camera: Keestoring Credibility to the photographic image[C]. IEEE Trans Consumer Electronics, 1993.905-910.
- [6] Clossn M, Gartshore S. Fully dynamic 3-dimensional orthogonal graph drawing[J]. Journal of Graph Algorithms and Applications, 2000, 5(2):1-34.