

doi: 10.3969/j.issn.1006-1576.2009.10.017

## 基于大背景的视频监控对 H.264 运动估计 JVT 推荐算法的改进

陈冒银<sup>1</sup>, 赵艳朝<sup>2</sup>

(1. 装备指挥技术学院 重点实验室, 北京 01416;

2. 中国人民解放军 63751 部队 测控与通信室, 陕西 西安 710000)

**摘要:** 针对大背景监控的视频内容特性, 对 JVT 在 JM10.2 中推荐的运动估计算法进行改进。利用运动矢量时空上的相关性, 对标准推荐算法的低精度估计部分进行裁减, 结合中心偏移假设和差值曲面单调性假设对高精度估计算法进行初步改进。测试数据证明, 对于大运动较少或准静止的视频, 改进后的运动估计快速算法在使得视频质量几乎不降的情况下, 将运算速度提高了 30% 左右。

**关键词:** 视频编码; H.264; 大背景监控; 运动估计

**中图分类号:** TP301.6 **文献标识码:** A

### Modification of Large-Background Video Surveillance on Recommended Arithmetic of H.264-ME-JVT

CHEN Mao-yin<sup>1</sup>, ZHAO Yan-zhao<sup>2</sup>

(1. Key Laboratory, Institute of Command & Technology of Equipment, Beijing 101416, China;

2. Team of Control & Communication, No. 63751 Unit of PLA, Xi'an 710000, China)

**Abstract:** According to video content characters of large-background surveillance application, improve the Motion Estimation algorithm recommended by JVT in JM10.2. Using correlations among motion vector space-time, and reduce the low accuracy evaluation part of the recommended algorithm. Through combining the two hypothesis of center excursion and monotonicity of differential surface, first improve the high-precision estimation algorithm. Test date proved that, with the same quality, the improved motion estimation algorithm has accelerated 30% than the old algorithm.

**Keywords:** Video encode; H.264; Large-Background surveillance; Motion estimation

## 0 引言

如何提高 H.264 运动估计技术的运算速度一直是视频编码技术研究的热点。通过研究各种应用场合视频信号的特点, 进行有针对性的运动估计算法改进是重要的研究方向。通常情况下, 大背景视频监控属于准禁止类视频, 即运动物体尺寸相对于背景比例小且运动幅度相对于背景尺寸也很小, 根据这一特性, 把 JVT 推荐的综合性能最优的算法进行裁减与改进, 在几乎不影响质量的前提下提高其运算速度, 使之成为适合大背景视频监控这一特殊应用场合的 H.264 快速运动估计算法。

## 1 JVT 推荐的运动估计快速算法

为保留不同研发厂商对编码器研发过程的自由度, H.264 标准并没有硬性规定采用某种具体的快速运动估计算法。但是在标准发展过程中, JVT 组织在对众多的快速算法提案进行比较分析后, 推荐了其中由清华大学的 Zhibo Chen, Peng Zhou, Yun He 等人提出的综合性能最优的算法<sup>[1-3]</sup>, 并在标准的参考代码 JM 中实现。该算法在搜索质量和速度

上都有理想的表现, 且对不同特性视频内容的适应性也非常好。整个算法把运动估计分为 2 部分, 即整数像素的运动估计及小数像素的运动估计。对于整数像素的运动估计采用非对称十字与复合六角型搜索算法 (Unsymmetrical-cross Multi-Hexagon-grid Search, UMHexagonS); 对于小数像素的运动估计算法则采用中心偏移小数像素搜索算法 (Center-Biased Fractional Pel Search, CBFPS)。

如图 1, 非对称十字与复合六角型搜索算法首先根据当前分块附近分块的运动矢量, 或者根据参考帧中对应分块的运动矢量, 对搜索起点进行粗略的预测, 然后以搜索起点为中心, 进行非对称十字搜索 (Unsymmetrical-cross Search, 如图中圆形像素点)。为适应视频内容中出现的大运动成分, 进行非均匀复合六角形搜索 (Uneven multi-hexagon-grid Search, 如图中正方形像素点)。经上述 2 个主要步骤的运动矢量搜索后, 无论视频内容中是大运动成分为主还是小运动成分为主, 都会把搜索得到的运动矢量初步推向最优运动矢量的附近范围。在此基础上, 整数像素估计算法的最后一个搜索步

骤，是以增强型的六角形搜索算法（Extended hexagon based Search，如图中菱形像素点）对运动

矢量进行细致的搜索，得出整数像素精度的最优运动矢量（3，-8）。

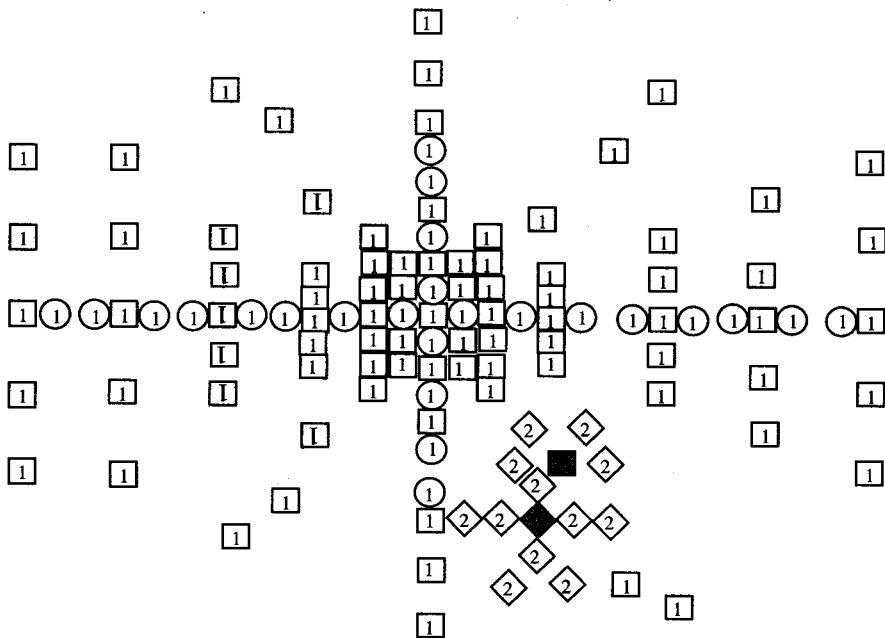


图1 UMHexagonS 算法示意图

为了获得更准确的运动矢量，提高视频数据的压缩比，推荐算法在完成整数像素精度的运动矢量搜索后，再进行小数像素精度的运动矢量搜索。小数像素精度的运动矢量搜索使用 CBFPS 算法：先利用整数像素精度运动矢量搜索前预测的结果与搜索得到的运动矢量结果进行预测，得到搜索起点，然后，采用 HEXBS 算法中的小六角形搜索模式（图 2，其中实心黑点表示整数像素所在位置）进行搜索，最后获得小数像素精度的最优运动矢量。

索算法（CBFPS）则是在此基础上得到更精确的最优运动矢量，属于高精度搜索。

在理解运动估计算法基本原理和标准推荐的快速运动估计算法后，可发现推荐算法力求兼顾实时和非实时性应用场合及各种特性的视频。而大背景视频监控的应用实时性要求比较高且图像背景属于准静止视频，针对该特性，分别对低精度的运动矢量搜索和高精度的运动矢量搜索进行改进。

## 2 运动矢量低精度搜索算法的裁减

在视频序列中，相邻的帧间常具有较高的相关性，在帧率较高时相关性更明显。同时，视频对象的运动具有连续的运动特征，故描述视频序列运动特征的运动矢量间也存在时间和空间上的相关性。

如图 3，设 A 为当前要编码的块，L、U、R、D 分别为其左侧、上侧、右侧和下侧的相邻块，A'、L'、U'、R'、D' 分别为前一帧中相对应的块，设 L、U、A、R、D 的运动矢量分别为  $V_1 \sim V_5$ ，各运动矢量都满足下式：

$$|V_i - V_3| \leq L (i=1,2,4,5)$$

则认为 A 块与相邻各块的运动矢量之间时空相关性较高，表明该区域的变化比较平缓，其中 L 为给定的阈值。否则，认为运动矢量之间相关性较低，表明该区域变化比较剧烈。

在推荐算法中，首先对大运动成分进行低精度

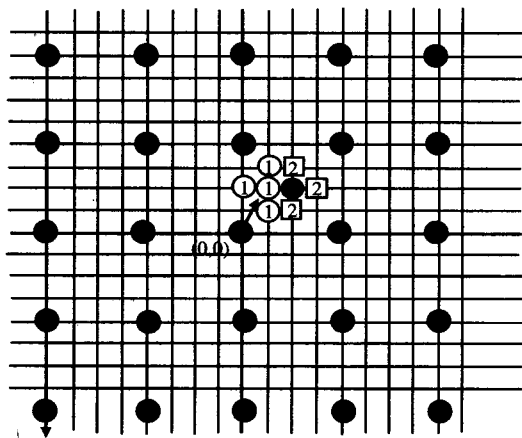


图2 CBFPS 算法示意图

可见，非对称十字与复合六角型搜索算法是为了将得到的运动矢量初步推向最优运动矢量所在的附近范围，属于低精度搜索。而增强型的六角形整数像素搜索算法（EHBS）和中心偏移小数像素搜

搜索, 搜索起点是利用当前块附近分块的运动矢量或者参考帧中对应分块的运动矢量进行匹配比较预测得到的。而根据实际应用的特点, 大背景视频监控的视频内容绝大部分是静止图像或准静止图像, 大运动成分极少, 其低精度搜索容易对系统资源造成浪费。根据以上分析, 如果当前块周边运动矢量满足式  $|V_i - V_3| \leq L (i=1,2,4,5)$ , 则当前分块附近区域运动矢量的时空相关性较高, 表明该区域变化比较平缓, 根据其预测得到的当前块的搜索起点, 就会非常接近最优运动矢量的位置。这种情况下, 无需进行非对称十字和非均匀复合六角形的低精度搜索, 可利用该搜索起点直接进入高精度搜索阶段以提高搜索速度。 $L$  的取值原则应是保证视频质量的前提下尽量提高搜索速度。如  $L=3$ , 可理解为当前分块周边运动矢量或参考帧中的周边运动矢量, 与参考帧中对应分块运动矢量之间的距离都小于 3 时, 只需要高精度搜索就能找到最优运动矢量。

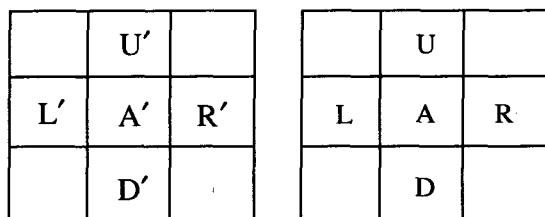


图3 运动矢量的时空域相关特性

### 3 运动矢量高精度搜索算法的改进

H.264 所有快速算法都是基于一定的理论假设对搜索模式进行不同程度的简化, 以达到在基本不影响搜索质量的情况下大幅提升搜索速度的目的。首先, 必须理解相关的理论假设, 其次, 再进行具体实现。从 BBGDS 算法后, 不限搜索步数的算法所依据的假设理论主要有 2 个: 一是中心偏移假设, 二是差值曲面的单调性假设。

中心偏移假设<sup>[4]</sup>指在自然视频中, 绝大部分的最优运动矢量都分布于搜索中心的附近范围, 即大运动矢量不占主导。这个假设通过对大量来自真实世界的视频内容进行统计分析得出, 而且这个假设尤其适合于类似大背景视频监控这类准静止的视频内容。该假设在算法的实现当中主要反映在算法的搜索步长上。如 BBGDS 算法每进行一步, 搜索仅向某个方向前进一个搜索点, 是一种搜索步长很紧密的策略, 这是因为在中心偏移假设的支持下, 用小步长就能取得很好的搜索效果。

差值曲面的单调性假设<sup>[5]</sup>指在矢量搜索过程

中, 每个搜索点所属的图像分块, 都会与参考帧中的指定分块根据匹配准则进行一次差值运算, 如果将搜索窗内各搜索点与其对应的差值运算结果关联起来, 就构成一个差值曲面。差值曲面的单调性假设认为在最优矢量附近, 差值曲面是单调的, 即搜索点越接近最优矢量, 就对应着越小的差值。该假设在算法的实现中主要反映在算法采用的搜索模式上: 围绕搜索起点和其周围等距离的若干搜索点进行搜索比较, 搜索起点随着算法的推进不断地接近最优矢量, 直到某步的搜索起点拥有比周围搜索点更小的差值, 满足了差值曲面的单调性假设, 搜索即可停止。之所以要与搜索起点等距离, 是因为在严格的差值曲面的单调性假设下, 这些点彼此存在任何优势, 或者说应拥有相同概率成为下步搜索起点。目前的快速算法中, 六角形的搜索模式更为接近于圆形, 在每步增加新搜索点的时候存在的冗余信息成分较少, 使之成为当今应用较多的算法。

标准推荐的增强六角形算法先进行粗略搜索 (大六角形搜索模式的半径为  $\sqrt{5}$ ), 最后细化搜索 (小十字形搜索模式的半径为 1), 所以对大运动成分较多的视频内容有很好效果, 但如果针对大背景监视这类准静止视频内容, 效率有提高的空间。

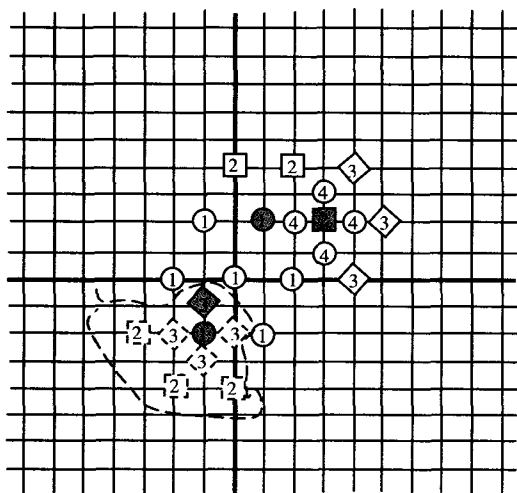
以上述理论为基础设计出的高精度改进算法把搜索分成 2 个子步完成, 且无论整数像素还是小数像素部分都采用同算法。2 个子步轮流进行直到算法满足条件而中止, 从而完成运动矢量的搜索过程。图 4 表示算法改进前后对运动矢量  $(-1, -1)$ ,  $(3, 2)$  搜索的过程, 其中虚线部分为  $(-1, -1)$  的搜索过程, 实线部分为  $(3, 2)$  的搜索过程。左图为改进前示意图, 右图为改进后示意图。以整数像素的改进为例, 改进后的算法过程如图 4。

算法第一子步: 对搜索起点及外围水平和垂直方向的 4 点搜索运算, 如果 4 个外围搜索点中有任一点对应的差值比搜索起点对应的差值小, 那么该点将作为下一个子步骤的差值最小点 (如图 4 的右子图中栅格点 1); 否则位于中间的搜索起点为最优的运动矢量所在, 结束搜索过程。理论依据是:

1) 大背景监视应用的准静止视频内容非常符合中心偏移假设, 可以采用搜索步长最短 (为 1) 的小十字搜索模式。

2) 小十字搜索模式相对于大六角形搜索模式更符合差值曲面的单调性假设, 因为大六角形模式中外围 6 个点距中心点距离并不是恒为  $\sqrt{5}$  (垂直

或水平方向为 2)，所以采用步长为 1 的小十字搜



索模式进一步减少了新增点的冗余成分。

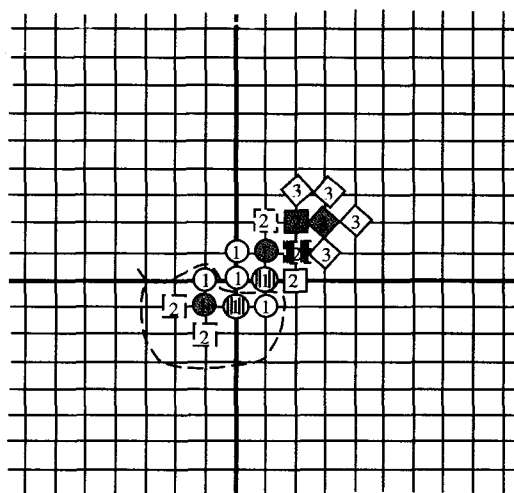


图 4 高精度算法改进前后示意图

算法第二子步：如果在完成第一子步骤的搜索后算法并没中止，则对距该大步起始点为  $\sqrt{2}$ ，且靠近栅格点 1 一侧的 2 点与栅格点 1 匹配运算，得到的差值最小的点将作为整个算法下一大步的搜索起点（如图 4 的右子图中阴影点 1）。理论依据：

1) 小十字算法虽然严格符合中心偏移假设和差值曲面的单调性假设，但是其每一步的新增点只有 3 个点，而改进后 2 个子步加起来新增点将达到 5 个，这样会提高搜索过程的收敛速度。

2) 第二子步新增的搜索点与起点的距离同为  $\sqrt{2}$  符合差值曲面的单调性假设。其它 2 个与起点的距离同为  $\sqrt{2}$  的点不参与匹配，因为它们距离差值最小点（栅格点 1）为  $\sqrt{5}$ ，而新增点距栅格点 1 为 1，根据差值曲面的单调性假设，新增点成为下一步搜索起点的优势极为明显，概率更高，舍弃其他两点减少了新增点的冗余成分。

通过对比图 4 中两算法示意图可见：

1) 如果起始点就是最优点所在，算法改进后，搜索点数由改进前的 11 点变为 5 点，减少了 6 个点，搜索效率提高 50% 以上。

2) 如果运动矢量为 (-1, -1)，算法改进后，搜索点数由改进前的 14 点变为 9 点，减少了 5 个点，效率提高 30% 以上。

3) 如果运动矢量为 (3, 2)，算法改进后，搜索点数由改进前的 17 点变为 16 点，减少了 1 个点，提高效率可忽略不计。

分析知，改进后的算法对运动矢量较小的块匹配运动估计，搜索效率将有很大的提升，特别适用于视频内容为准静止状态的大背景视频监控应用。

#### 4 改进后的运动估计算法

在分析原代码的基础上，对代码重新实现，从而在不影响整体编码流程下，实现改进后的运动估计算法。JVT 小组推出的软件参考模型 JM10.2 中，负责实现非对称十字与复合六角型搜索算法的函数是 FastIntegerPelBlockMotionSearch() 函数，负责实现中心偏移小数像素搜索算法(CBFPS)的函数是 FastSubPelBlockMotionSearch() 函数。由于对整数像素高精度运动估计的算法改进和小数像素高精度运动估计的算法改进一致，对于小数像素运动估计部分的算法改进可参考整数像素高精度部分，具体分析可结合代码中的注释理解。

由于 JM 代码本身是为算法研究及标准的发展提供统一的平台，故除了对 H.264 标准的功能作全面实现外，还提供各种实验研究所需的测试参数。对于本部分研究的运动估计模块，也提供了专门的时间统计数据，现将通过对 2 个类似大背景视频监控应用的测试序列(silent\_qcif.yuv, akiyo\_qcif.yuv)测试（测试序列长度为 100 帧），实验平台：奔腾(R) D CPU 2.8G 2.8G 512 M 内存，WindowsXP。表 1 给出了具体测试参数。

表 1 运动估计算法改进前后测试参数对比

序列	指标	每帧时间 (s)	每帧运动估计时间 (s)	SNRY (dB)
	算法			
silent_qcif.yuv	标准推荐算法	1.579	1.008	35.88
	改进后算法	1.353	0.775	35.82
akiyo_qcif.yuv	标准推荐算法	1.330	0.810	38.24
	改进后算法	1.073	0.535	38.19
均值	标准推荐算法	1.455	0.909	37.06
	改进后算法	1.213	0.655	37.01

$$SCR = SCR\_base * 300 + SCR\_ext$$

$$SCR\_base = ((system\_clock\_frequent * t) / 300) \% 2^{33}$$

$$SCR\_ext = ((system\_clock\_frequent * t) / 1) \% 300$$

如果记录时间小于 26.5 h, SRC 可以由下面公式与 10 MHz 的 RTC 相互转换:

$$10 \text{ MHz RTC time} = SRC * 10 / 27 \text{ (取整数)}$$

如果记录时间超过了 25.6 h, 那么格式 0 视频包头时间标签可以用来求出 MPEG2 的 SCR 翻转次数和 RTC 的高 8 位的值。

此 system\_clock\_frequent 为 27 MHz 的参考时钟。因 IRIG106 提供的这 10 MHz 的 RTC 为高精度的时钟, 其精度可通过时间补偿及同步于 GPS 时钟信号来增强, 故通过 RTC 生成的 MPEG2 的 27 MHz

\*\*\*\*\*

(上接第 56 页)

通过表 1 中数据可以得出 2 条结论:

1) 从均值上看, 算法改进后, 每帧编码时间相对于推荐算法减少了 16%, 每帧运动估计时间减少了 28%, 视频质量下降 0.05 dB。

2) 改进后算法对不同特征视频的作用效果不同: 对于大运动成分较多的 silent\_qcif.yuv 序列, 由于算法的改进每帧运动估计时间减少了 13%, 视频质量下降 0.06 dB; 对于大运动成分较少的 akiyo\_qcif.yuv 序列, 每帧运动估计时间减少了 34%, 视频质量下降 0.05 dB。表明算法的改进对于大运动较少或准静止的视频效果更佳。

## 5 结束语

测试数据可见: 把 JVT 推荐的算法分为高精度和低精度, 结合大背景监控视频的特性分别裁减与改进, 在原视频质量略有下降下, 提高了 H.264 运动估计算法的计算速度。该算法对越接近静止类的视频应用效果越好, 适合大背景视频监控应用场合。

## 参考文献:

- [1] Zhibo Chen, Yun He. Fast Integer and Fractional Pel Motion Estimation [J]. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG 5<sup>th</sup> Meeting: Geneva, Switzerland, JVT-E045, 2002(10): 9-17.
- [2] Zhibo Chen, Peng Zhou, Yun He. Fast Integer Pel and Fraction Pel Motion Estimation for JVT [J]. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG 6<sup>th</sup> Meeting: Awaji, Island, JP, 2002(12): 5-13, JVT-F017.
- [3] Zhibo Chen, Peng Zhou, Yun He. Fast Motion Estimation for JVT[J]. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG 7<sup>th</sup> Meeting: Pattay II Thailand, 7-14 March, 2003, JVT-G016.
- [4] Lurng-Kuo Liu, Ephraim Feig. A Block-Based Gradient

精确度也增高, 这对分布式数据的同步至关重要。

## 5 结束语

跟踪 IRIG106 标准, 根据客户需求研发了符合 IRIG106 第十章标准的音视频压缩板卡、MIL1553 数据采集板卡、1394B 总线数据存储等板卡, 用于嵌入式实时数据存储系统中, 取得好的应用效果。

## 参考文献:

- [1] Richard Bond. The potential of digital flight data recording [C/OL]. Electronics talk Editorial Team on 2007, 24(1).
- [2] RCC Document 106-07. Telemetry Standard [S]. September 2007(9).
- [3] International Standard ISO/IEC 13818. Information Technology Generic Coding of Moving Pictures [S]. 2000(2).

Descent Search Algorithm for Block Motion Estimation in Video Coding[J]. IEEE Transactions on Circuits And Systems for Video Technology, 1996(8): 419-422.

- [5] Shan Zhu, Kai-Kuang Ma. A New Diamond Search Algorithm for Fast Block Matching Motion Estimation[D]. Singapore: International Conference on Information, Communications and Signal Processing ICICS, 1997: 292-296.
- [6] 白然施. 视频压缩中的运动估计技术的研究[D]. 汕头: 汕头大学硕士学位论文, 2003: 15.

\*\*\*\*\*

(上接第 58 页)

野外指挥自动化网络的电源保障是网络系统的基础, 供电系统的好坏直接影响指挥自动化系统的正常运行。野外地域提供的电源一般存在电压不稳、变化范围大、电源干扰严重、没有接地等问题, 故在网络设备加电运转前, 应先将电源接入稳压器和 UPS 电源, 并将网络设备按规范严格接地。

## 4 结束语

野外指挥自动化网络系统是复杂系统。通过考虑管理与技术, 制定管理条款及安全保密策略, 并采用数据加密及访问控制等手段, 以快速准确地建立稳定可靠、安全保密的野外指挥自动化网络系统。

## 参考文献:

- [1] 许金裕. 21 世纪信息战理论研究及探索[M]. 北京: 解放军出版社, 2005.
- [2] 王丽, 姜建安, 赵维霞. 指挥自动化网络安全系统的构建[J]. 兵工自动化, 2006, 25(3): 26-27.
- [3] 陈利明, 邵芳. 浅析网络信息安全的综合保护[J]. 兵工自动化, 2008, 27(8): 45-46.
- [4] 单连平, 王亚, 吴维元. 指挥自动化系统现状与发展趋势[J]. 指挥控制与仿真, 2007(3): 1-5.
- [5] 沈兆吉. 关于军队指挥自动化建设的几个问题[J]. 军事系统工程, 1999(3): 5-7.