基于四元数奇异值的数字视频质量客观评价方法应用研究

王 佳，司占军
（天津科技大学，天津 300222）

摘要： 针对 512 kb/s，1 Mb/s，2 Mb/s 这 3 种中国常见的网络环境，以及 FLV，WMV，MP4 这 3 种国内流行的流媒体格式压缩视频，利用四元数奇异值法，对同一视频压缩成的不同格式、码率、分辨率的数字视频进行客观评价，并利用主观实验验证该方法的准确性。评价结果显示：512 kb/s 下，FLV 格式以不错的清晰度得到了较好的评价；1 Mb/s 和 2 Mb/s 下，WMV 格式较 FLV 和 MP4 格式的评价较差；2 Mb/s 以上的带宽，MP4 格式视频失真程度最小，因此在清晰度和观影视觉效果方面具有非常大的优势。

关键词： 流媒体；数字视频客观评价方法；四元数奇异值法
中图分类号： TS807 **文献标识码：** A **文章编号：** 1001-3563(2012)07-0112-06

Application Study of Digital Video Quality Objective Evaluation Method Based on Quaternion Singular Value

WANG Jia, *SI Zhan-jun*
(Tianjin University of Science & Technology, Tianjin 300222, China)

Abstract: Aiming at three common Internet environments in China and three popular formats as FLV, WMV, and MP4, objective evaluation method of Quaternion singular value was applied to evaluate digital videos with different format, code rate, resolution that are compressed from the same video, and using subjective experiment to validate its accuracy. The result showed that FLV format gets better evaluation with the good resolution under 512 kb/s; WMV gets lower evaluation than FLV and MP4 under 1 Mb/s and 2 Mb/s; Over 2 Mb/s, video of MP4 format has a lowest distortion, so it has obvious advantage on resolution and visual effect.

Key words: streaming media; digital video objective evaluation method; quaternion singular value

近几年，流媒体技术飞速发展，并将在未来为全世界创造惊人的财富[1]。流媒体领域在不断壮大，从最初的音频扩展到视频，但是由于网络等原因，一直无法对流媒体进行准确评价。对于数字视频客观评价方法的研究，已经有了诸多的成果，但对于能够在任何带宽下活跃的流媒体来说，如何在有限的带宽下得到质量较好的数字视频，一直存在争议。本文主要研究 FLV，MP4 和 WMV 3 种格式的视频[2]在流媒体 512 kb/s，1 Mb/s，2 Mb/s 带宽下的表现，通过客观评价四元数奇异值法得到 3 种带宽下每种格式的分数，进而得出视频格式与网络带宽对应的关系，为客户带来更直接、更美的享受[3]。

1 四元数奇异值算法介绍

四元数奇异值法因四元数的构建及色彩空间的不同而不同。该算法采用 YCbCr 空间，四元数选择像素的独立性质——亮度和色度，以及反映空间相关性和时间相关性的边缘能量和残差能量，并将人眼最敏感的亮度作为实部。四元数奇异值算法具有非常好的优势，它既保持了自己的客观性，又和人眼的相关性非常大，能较正确地和人眼所观看的结果取得一致，是目前较新的客观评价算法[4]。

四元数模型[4]如公式(1)所示：

$$S(x,y)=Y(x,y)+C(x,y)\mathrm{i}+E(x,y)\mathrm{j}+R(x,y)\mathrm{k} \tag{1}$$

收稿日期： 2011-11-07
作者简介： 王佳(1987－)，女，内蒙古包头人，天津科技大学硕士生，主攻印刷领域数字出版。

其中：Y 表示该像素的亮度，作为实部；C 表示该像素的色度，作为第一个虚部，色度通过 YCbCr 空间和 RGB 空间之间的转换获得；E 使用了 sobel 算子进行边缘算法计算，在这里使用的 sobel 算子是{{−1,0,1},{−2,0,2},{−1,0,1}}和{{−1,−2,1},{0,0,0,},{1,2,1}}，对于图像的每一个像素 $P(x,y)$，通过 sobel 算子计算它的“梯度向量信息”（$G_x(x,y)$，$G_y(x,y)$），其中 $G_x(x,y)$和 $G_y(x,y)$分别为垂直边缘算子和水平边缘算子，定义像素点 $P(x,y)$的边缘能量 $E(x,y)$为：

$$E(x,y)=\sqrt{[G_x(x,y)]^2+9\times[C_y(x,y)]^2} \quad (2)$$

$E(x,y)$作为第 2 个虚部；R 作为第 3 个虚部，表示“残差能量”，本实验定义了第 T 帧像素点 $P(x,y)$相对于其前 i 帧（即 $T-i$ 帧）的残差能量 $R(x,y)$为：

$$R(x,y)=Y_T(x,y)+C_T(x,y)-Y_{T-i}(x,y)-C_{T-i}(x,y) \quad (3)$$

i 为间隔差，当 i 为 1 时，如果不是场景切换，残差图像绝大部分位置的残差能量为 0，对于两帧图像间的时间相关性反映不是很突出，所以设定残差间隔 i 为 3。

四元数各部分的具体计算及实现参照程序编写部分。为了便于计算，四元数建立完毕后，按照 8×8 来分隔。通过奇异值分解（SVD）[5]，每个四元数矩阵都得到一个奇异值向量，且向量的每个元素都是实数，因此本文利用原始和参考块的奇异值向量的欧式距离，来度量对应的块失真[4]：

$$D_i=\sqrt{\sum_{j=1}^{n}(s_j-\hat{s}_j)^2} \quad (4)$$

其中：s_j 和 $\hat{s}_j$ 分别为参考和失真块对应的奇异值；n 为块的大小。结果数值越大，表明差异越明显，失真越严重。

2　实验

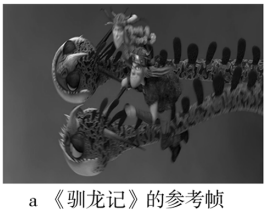
2.1　准备视频

为了使客户可以根据视频类型选择带宽及格式观看，根据日常观看的电影类别选择了以下 3 个较有代表性的视频进行实验：动画电影《驯龙记》、人物电影《Big. Fish》和风景介绍《日本之美》。《驯龙记》是一部计算机制作的动画电影，拥有优秀的画面和庞大的数据量；《Big. Fish》是一部人物特写比较多的电影，数据量比较小；《日本之美》是自然风景电影，数据量根据画面细节而定，描写植物时需要的颜色种类多，数据量大，建筑物则由于颜色单一，需要的数据量小。此外，要想压缩的结果视频具有良好的效果，需要高清视频（HD）。高清视频可提供相对于传统模拟电视技术更高清晰度的图像质量，早在 2006 年 4 月，国家信息产业部就发布了高清视频的标准，要求清晰度达到 720 线以上且画面比例必须是 16∶9。所以需要的高清视频标准应该是 1 280×720（720 P）或 1 920×1 080（1 080 P）的分辨率。目前，获取高清视频的途径主要是购买蓝光光盘和从网络下载[6]。本文所用到的 3 个视频均是从网络上下载的高清视频。

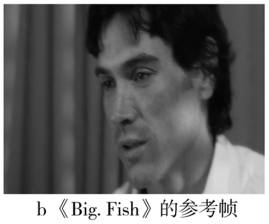
2.2　压缩视频并截取实验帧

所用到的压缩软件是 XMedia Recode，是国外人士对各个公司发布的压缩工具的整合，基本保留了原始压缩的信息和所有的可调整参数，由于它不需要任何插件，所有的压缩方法都集成到程序里面，也不需要额外安装，给用户带来了方便。本实验所用到的截帧软件是 Adobe Media Encoder CS4 和 KmPlayer。

根据流媒体格式，将视频压缩为 MP4，WMV，FLV 3 种格式，因其封装和格式不同，因此步骤和参数略有差别。对应 512 kb/s，1 Mb/s，2 Mb/s 的带宽码率分别为 600 kb/s，960 kb/s，1 600 kb/s[7]，分辨率为 640×480（4∶3）和 1 280×720（16∶9）。分别截取压缩前后视频的同一帧以及压缩后试验帧的后三帧进行评价。由于不同画面不同帧的四元数奇异值差异较大，无法取平均值保证准确性，因此选取帧时要注意所选试验帧能够最佳体现视频类型，并保证相同视频截取的试验帧为同一帧，以确保实验准确性。分辨率 640×480 下，3 个实验视频压缩前的参考帧见图 1。



a《驯龙记》的参考帧



b《Big. Fish》的参考帧

c《日本之美》的参考帧

图 1　3 个视频压缩前的参考帧

Fig. 1 The reference frame of three videos before compression

2.3 程序编写

实验算法由 C# 编写,具体如下:

取得亮度和色度的代码:

```
public void rgbre(string q)
{
Bitmap w = new Bitmap(q);
width = w.Width; height = w.Height;
r1 = new int[width, height]; g1 = new int[width, height]; b1 = new int[width, height]; y1 = new int[width, height]; br1 = new int[width, height];
Color pixtel1;
for (int i = 0; i < width; i++)
{
for (int j = 0; j < height; j++)
{
pixtel1 = w.GetPixel(i, j);
r1[i, j] = pixtel1.R; g1[i, j] = pixtel1.G; b1[i, j] = pixtel1.B;
y1[i, j] = (int)(0.399 * r1[i, j] + 0.587 * g1[i, j] + 0.114 * b1[i, j]);
br1[i, j] = (int)(0.564 * (b1[i, j] - y1[i, j]));
}
}
}
```

Sobel 算子的代码:

```
public void so(int[,] r,int[,] g,int[,] b,int a, int s)
{
v9 = new int[a, s]; v11=new int[a,s]; v22= new int[a,s];
for(int q=0;q<a-2;q++)
for(int w=0;w<s-2;w++)
{
v0 = (int)(r[q , w ] * 0.3 + 0.59 * g[q , w ] + 0.11 * b[q , w ]);
v1 = (int)(r[q+1, w ] * 0.3 + 0.59 * g[q +1, w ] + 0.11 * b[q+1, w ]);
v2 = (int)(r[q +2, w ] * 0.3 + 0.59 * g[q +2, w ] + 0.11 * b[q+2 , w ]);
v3 = (int)(r[q , w+1] * 0.3 + 0.59 * g[q, w+1] + 0.11 * b[q, w+1]);
v4 = (int)(r[q+1, w+1 ] * 0.3 + 0.59 * g[q +1, w +1] + 0.11 * b[q+1, w+1]);
v5 = (int)(r[q+2, w+1] * 0.3 + 0.59 * g[q+2, w+1] + 0.11 * b[q+2, w+1]);
v6 = (int)(r[q, w+2] * 0.3 + 0.59 * g[q, w+2] + 0.11 * b[q, w+2]);
v7 = (int)(r[q+1, w+2] * 0.3 + 0.59 * g[q+1, w+2] + 0.11 * b[q+1, w+2]);
v8=(int)(r[q+2, w+2 ] * 0.3 + 0.59 * g[q+2 , w+2 ] + 0.11 * b[q+2, w+2]);
v11[q, w] = -1 * v0 + v2 - 2 * v3 + 2 * v5 - v6 + v8;
v22[q, w] = -1 * v0 - 2 * v1 - v2 + v6 + 2 * v7 + v8;
v9[q, w] = (int)(Math.Sqrt(v11[q, w] * v11[q, w] + v22[q, w] * v22[q, w]));
}
}
```

程序窗口界面见图 2,利用此窗口,选择已截取

图 2 程序窗口界面

Fig. 2 Program window interface

好的试验帧和参考帧,即可计算出相应的四元数奇异值的结果分值。分值越大,表明失真越严重。

2.4 评价标准

根据五级制主观评价标准[8],给出四元数奇异值算法的评价标准,见表 2。

表 2 四元数奇异值算法的评价标准

Tab. 2 The evaluation standard of quaternion singular value

级数	分值	评价标准
1	20 分以下	和参考帧几乎一样,能非常好地享受
2	20～50 分	能感觉到失真,但是不影响观看
3	50～110 分	失真比较明显,稍微不能忍受
4	110～150 分	失真非常明显,不能忍受
5	150～200 分	无法观看,令人讨厌

3　结果与讨论

本实验分别给出 3 种压缩格式下，分辨率为 640×480 和 1 280×720 时，不同码率、不同风格视频的四元数奇异值结果。

3.1　MP4 格式

MP4 格式是国际通用格式，近几年在 H264 编码的支持下，MP4 格式的视频在流媒体方面取得了非常大的优势。对 3 部影片码率从 600 kb/s 到 1 600 kb/s 进行评价，在分辨率 640×480 和分辨率 1 280×720 下客观评价的四元数奇异值结果分值分别见图 3。

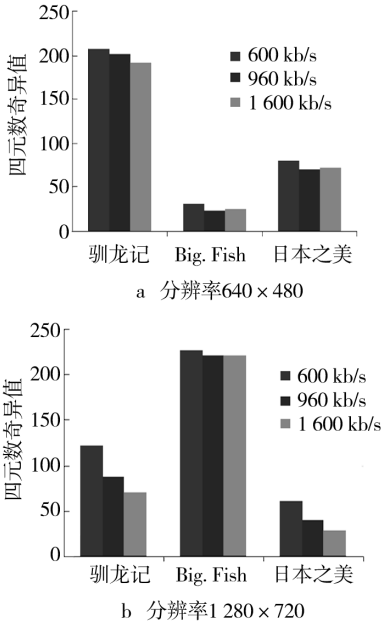


图 3　MP4 格式视频四元数奇异值的变化

Fig. 3 The variation of quaternion singular value under MP4 format

3.2　WMV 格式

WMV 格式是微软制作的一款视频格式，在低码率下表现情况不佳，此格式适合保存本地视频，不适合进行传播。在分辨率 640×480 和分辨率 1 280×720 下客观评价的四元数奇异值结果分值见图 4。

3.3　FLV 格式

FLV 在几年前凭借 H. 264 编码，风靡全球一段非常长的时间，目前由于网络速度的提高和手机网络的发展，FLV 文件退居二线，但是 FLV 在低码率的表现仍旧十分出色。在分辨率 640×480 和分辨率

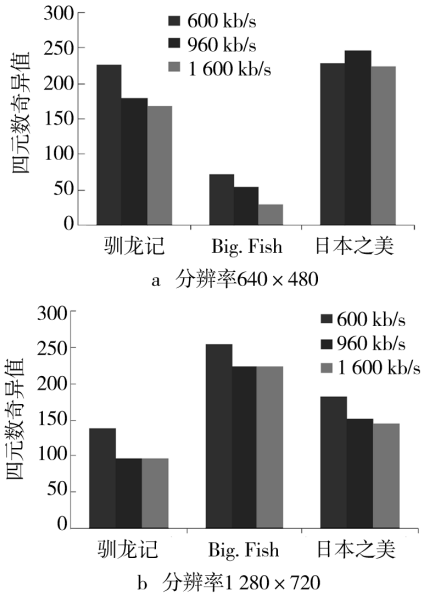


图 4　WMV 格式视频四元数奇异值的变化

Fig. 4 The variation of quaternion singular value under WMV format

1 280×720 下客观评价的四元数奇异值结果分值见图 5。

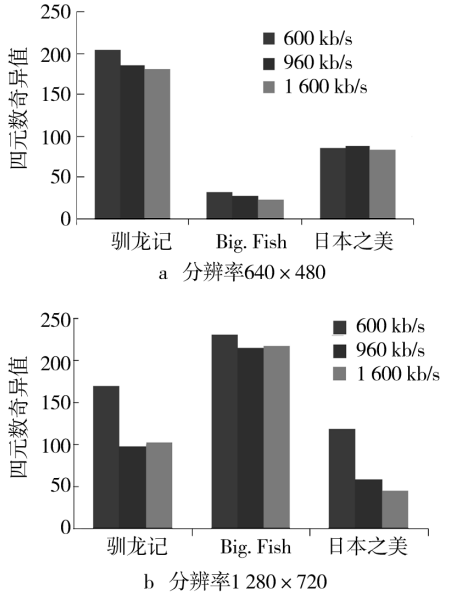


图 5　FLV 格式视频四元数奇异值的变化

Fig. 5 The variation of quaternion singular value under FLV format

3.4　结果分析

1）分辨率为 1 280×720 时，无论何种码率和格式，人物电影《Big. Fish》的四元数奇异值结果均在 200 分以上，属于无法观看的程度，但实际情况是此条件下

此视频仍然可以观看。初步分析这一特殊情况，是因为人物电影的人物特写较多，高分辨率下面部细节更为丰富，因此评价较差。从这一结论可以看出，四元数奇异值法用于评价人物脸部细节描写较多的视频时存在较大误差，对于此类视频的评价并不适用。

2）分辨率 1 280×720 下视频的得分比分辨率 640×480 的基本要低一些，而 600kb/s 码率下分数要低很多，这充分说明了分辨率与码率的正比关系，要想得到相同的效果，提高分辨率，码率也要相应提高。

3）1 280×720 分辨率下，推荐 MP4 格式。分辨率 640×480、低码率下，MP4 格式和 FLV 格式不相上下，但是 FLV 格式在文件大小上获得了优势。

4）分辨率高，在足够的数据量支持下，可以为画面增加很多细节，能提高整体的画面效果。高分辨率主要是为了电影的观赏，而能不能看高清和看怎样的高清，这个问题一直没有定论。根据以上实验结果，建议看高清视频时，先要了解电影类型。如果是数据量较大的视频，例如《驯龙记》这样的动画电影，推荐使用 1 Mb/s 以上的带宽和相应码率的视频；真人电影且画面激烈程度不高的，512 kb/s 的带宽就能得到清晰的效果，画面比较激烈的推荐 1 Mb/s 以上的带宽和相应码率；对于描写风景类的电影，由于描写天气和树木等细节，对数据量的需求较大，推荐 1 Mb/s 带宽及相应的码率。

5）512 kb/s 带宽推荐 FLV 格式，FLV 的文件所需数据量小，画质、清晰度等表现不错，但是由于带宽的限制，不推荐观看 3D 动画电影和激烈场景较多的动作电影。1 Mb/s 带宽推荐 MP4 格式。前面提到，WMV 格式在高码率下有非常好的得分，而在低码率下得分较差，通过码率的变化也可以看出，WMV 是一个对数据量敏感的格式，因此不推荐观看。

6）整体来说，1 Mb/s 以下带宽推荐 FLV 格式，1 Mb/s 以上推荐 MP4 格式。

7）类似《驯龙记》的动画电影和类似《Big. Fish》的人物电影，MP4 格式和 FLV 格式都可以选择。类似《日本之美》的风景电影推荐 FLV 格式。

8）MP4 格式会成为主导格式，活跃于移动流媒体和计算机流媒体。

3.5 基于心理物理学的主观实验验证

由于实验结论得出此客观评价方法并不适用于人物面部细节较多的电影，因此选择另外两部视频进行主观实验以验证结果。选取 10 名观察者观看分辨率 1 280×720，码率分别为 600，900 和 1 600 kb/s，格式为 MP4 的动画电影和风景电影，根据表 2 的 5 级制主观评价标准对视频打分，并根据表 2 将客观评价结果（四元数奇异值结果分值）换算为五分制，主、客观评价结果的一致性见图 6。

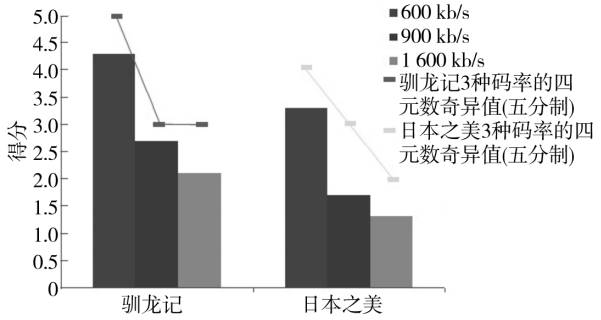


图 6 主客观评价结果的一致性

Fig. 6 The consistency between subjective evaluation and objective evaluation

由于换算客观评价结果时采用的五级制只有 5 个分值，没有更为细致的评分尺度，因此换算后的客观评价结果与主观评价结果存在 1 分以下的差别，但是从差别及结果趋势仍可以看出，主观评价结果与客观评价结果一致性较好，即四元数奇异值法与主观评价结果有很好的相关性，对于非高分辨率人物电影的视频评价基本准确。

4 结语

研究结果可以针对不同风格的视频，根据不同的需要，为用户提供更为优质的流媒体视频，给客户带来更美的视觉享受[9]。相信未来流媒体的应用范围会越来越广泛，成为未来世界媒体界的主导。

参考文献：

[1] 余永海. 从形式到内容的创新——谈数字出版物的设计开发[J]. 包装工程，2007，28(9)：168－170.
YU Yong-hai. Innovation in Digital Publications Design [J]. Packaging Engineering，2007，28(9)：168－170.

[2] 刘义良. 基于视频格式的转换与研究[J]. 计算机与数字工程，2010，247(5)：178－180.
LIU Yi-liang. Analysis and Research on Video Format Conversion[J]. Computer & Digital Engineering，2010，

247(5):178－180.

[3] 陈昭霓.视频媒体中的动态版式设计[J].包装工程,2010,31(3):99－102.
CHEN Shao-ni. Dynamic Typography Design in Video Media[J]. Packaging Engineering, 2010, 31(3): 99－102.

[4] 张富强.基于四元数奇异值分解的视频质量评价方法[J].电子学报,2011,39(1):219－223.
ZHANG Fu-qiang. Video Quality Assessment Based on Quaternion Singular Value Decomposition[J]. Acta Electronica Sinica,2011,39(1):219－223.

[5] 朱大鹏,周世生,张志昆.基于改进的 Prony 方法的包装系统参数识别[J].包装工程,2009,30(10):84－86.
ZHU Da-peng,ZHOU Shi-sheng,ZHANG Zhi-kun. Parameter Identification of Packaging System with Modified Prony Method[J]. Packaging Engineering,2009,30(10):84－86.

[6] 申健明.剪辑·特效·字幕·压制全攻略[M].北京:人民邮电出版社,2009.
SHEN Jian-ming. All Strategy of Montage, Special Effects,Subtitle,and Suppression[M]. Beijing:People's Post and Telecommunication Press,2009.

[7] 胡媛,司占军.压缩格式和码率对视频质量的影响及应用研究[J].中国印刷与包装研究,2011,3(3):21－26.
HU Yuan,SI Zhan-jun. Study on Effect of Compression Format and Code Rates on Video Quality and Application[J]. China Printing and Packaging Study, 2011, 3(3):21－26.

[8] 宋慧慧,赵秀萍,司占军.基于主观评价法对高光泽喷墨打印纸图像质量的研究[J].包装工程,2008,29(5):39－42.
SONG Hui-hui,ZHAO Xiu-ping,SI Zhan-jun. Research on Image Quality of Glossy Ink Jet Paper Based on Subjective Evaluation[J]. Packaging Engineering, 2008, 29(5):39－42.

[9] 尚晓明.信息时代背景下视觉传达设计的新语境[J].包装工程,2009,30(11):241－242.
SHANG Xiao-ming. New Context of Visual Communication Design under the Background of Information Age[J]. Packaging Engineering,2009,30(11):241－242.

(上接第 23 页)

[5] 孟婕,孙诚,王建清,等.包装印刷中稀土荧光防伪油墨的研究[C]//第十三届全国包装工程学术会议.2010:704－707.
MENG Jie, SUN Cheng, WANG Jian-qing. The Study of Rare-earth Fluorescent Falsification-resistant Ink in Packaging and Printing[C]//The 13th National Academic Conference of Packaging Engineering. 2010:704－707.

[6] 湛雪辉,赵学辉,湛含辉.苯甲酸及其衍生物和噻吩甲酰三氟丙酮铕配合物的合成与荧光性质[J].湖南师范大学自然科学学报,2008,31(1):82－84.
ZHAN Xue-hui,ZHAO Xue-hui,ZHAN Han-hui. Synthesis and Fluorescent Properties of the Complexes of Benzoic Acid or Its Derivatives with α-thenoyltrifluoroacetone Europium (Ⅲ) [J]. Journal of Natural Science of Hunan Normal University,2008,31(1):82－84.

[7] NAKAMURA Kazuki,HASEGAWA Yasuchika. Enhanced Lasing Properties of Dissymmetric Eu(Ⅲ) Complex with Bidentate Phosphine Ligands[J]. J Phys Chem A, 2007,111:3029－3037.

[8] 江祖成,蔡汝秀,张华山.稀土元素分析化学[M].第 2 版.北京:科学出版社,2000.
JIANG Zu-cheng,CAI Ru-xiu,ZHANG Hua-shan. Analytical Chemistry of Rare Earth Elements[M]. 2nd Edition. Beijing:Science Press,2000.

[9] 郝晓秀,杨淑蕙,孙诚.荧光防伪包装纸荧光强度的测定技术[J].包装工程,2007,28(9):24－26.
HAO Xiao-xiu, YANG Shu-hui, SUN Cheng. Fluorescent Intensity Determination Technology for Anti-counterfeit Packing Paper[J]. Packaging Engineering, 2007, 28(9):24－26.

(上接第 100 页)

[3] 李军,高德,卢富德.包装机中曲柄滑块机构给定两对相应位移的解析法设计[J].农机化研究,2008(7):111－112.
LI Jun,GAO De,LU Fu-de. The Design of the Analytical Method on Slider-crank Mechanism in Given the Two Related Displacements[J]. Journal of Agricultural Mechanization Research,2008(7):111－112.

[4] 陈杰,梁爱锋,吴淑芳.裹包机中的曲柄-摇杆机构优化设计[J].包装工程,2012,33(3):75－76.
CHEN Jie, LIANG Ai-feng, WU Shu-ang. Optimization Design of Crank-rocker Mechanism of Packaging Machine[J]. Packaging Engineering,2012,33(3):75－76.