

Video Transport Evaluation with H.264 Video Traces

Sha Bai, Jingwen Li, Xiaoshuai Cui, Yin Li, Junbo Gao

CNPC Well Logging Huabei Business Division, Renqiu Hebei
Email: baisha0110@163.com

Received: Nov. 2nd, 2017; accepted: Nov. 14th, 2017; published: Nov. 21st, 2017

Abstract

This tutorial introduces a trace-based evaluation methodology for the network transport of H.264 encoded video. We first give an overview of H.264 video coding, and then present the trace structures for capturing the characteristics of H.264 encoded video. We give an overview of the typical video traffic and quality characteristics of H.264 encoded video. Finally, we explain how to account for the H.264 specific coding mechanisms, such as hierarchical B frames.

Keywords

H.264 Encoded Video, Hierarchical B Frames, Medium Grain Scalability (MGS), Network Transport, Video Trace

支持H.264视频跟踪的视频传输评价

白 莎, 李静文, 崔小帅, 李 印, 高俊博

中国石油测井有限公司华北事业部, 河北 任丘
Email: baisha0110@163.com

收稿日期: 2017年11月2日; 录用日期: 2017年11月14日; 发布日期: 2017年11月21日

摘 要

论文介绍了一种基于H.264编码视频的网络传输跟踪评价方法。首先概述H.264视频编码; 然后介绍了捕捉H.264编码视频特点的跟踪结构, 给出一种典型的视频流量和H.264编码视频的质量特征; 最后, 解释了H.264特定编码机制, 如分层B帧。

关键词

H.264编码视频, 分层B帧, 媒体的可扩展性(MGS), 网络传输, 视频跟踪

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 简述

随着编码视频在网络流量中占比的不断增长, 视频传输机制的性能评估变得越来越重要。论文提出一个基于 H.264 视频编码标准机制的视频传输网络跟踪综合评价方法, 对视频网络传输的性能评估主要从两方面提出, 即增加视频网络流量和高效的 H.264 视频编码标准。论文总体目标是设计 H.264 编码视频的传输机制, 以及通过 H.264 视频追踪对 H.264 视频传输机制的性能进行评估。

2. H.264 视频编码概述

H.264 视频编码标准的增强块变换和帧间编码及运动补偿预测的帧内编码之间组合构成编码循环[1]。每个视频帧根据图片结构分为内部编码(I), 向前预测编码(P)和运动补偿预测 I 或 P 帧, 或双向预测编码(B)。通过经典的 B 帧预测, 一个 B 帧编码与运动补偿预测可以预测并实现 I 或 P 帧。经典 B 帧预测被用于预测 MPEG 标准并默认 H.264/AVC, 但是 B 帧不能作为预测另一个 B 帧的参考。增强编码机制 H.264 建立在基本的 MPEG 视频编码循环, 与早期 H.264 视频编码标准的效率相比, 作为一个累积效应得到显著改善。

2.1. 单层(非可扩展性) H.264 视频编码

1) H.264/AVC: H.264/AVC 标准介绍了可变尺寸的模块, 如 4×4 , 8×4 , 16×8 和 8×8 像素为提高基础模块的预测编码。帧内编码是通过空间帧内预测改进, 即从一个相同帧的类似模块预测一个给定的帧模块。帧间预测通过不平等的多重参考帧进行提高。

编码循环是通过减少引入的基础模块进行循环解封过滤器而得到进一步提高的编码。MPEG 视频编码循环执行块变换, 其次是量子化的变换系数, 以及随后的锯齿形扫描和游程长度编码。由此产生的符号通过熵编码被进一步压缩。在 H.264/AVC 中, 熵编码通过上下文自适应算术编码(CABAC)得到改进, 这是高效但计算密集型, 或上下文自适应可变长编码(CAVLC), 这是一个改进的前可变长度的编码机制。单独编码决定和整个编码循环通过拉格朗日进行优化改进, 旨在共同实现最小化比特率, 并使得引入的视觉失真最小化。

H.264/AVC 标准分类了所有的视频编码机制, 有所谓的视频编码层(VCL)。此外, H.264/AVC 标准定义了一个网络抽象层(NAL)包含视频编码数据映射到一个网络传输层的功能[2]。这个视频编码数据被组织成 NAL 单元(NALUs)。每个 NALU 包含一个整数字节的视频编码数据, 以及一个 1 字节的头。

2) 单层 H.264 SVC 编码: 虽然 H.264 SVC 主要介绍对 H.264/AVC 添加可伸缩的编码特性, H.264 SVC 有着重要的增强, 提高单层编码的功能。视频编码标准 H.264/AVC 前期严格遵循经典 B 帧预测, 如图 1(a)所示, 即 B 帧不能用来预测其他 B 帧。H.264/AVC 通过广义的 B 帧概念取消这一限制, 允许 B 帧块作为其他运动补偿的预测 B 帧的参考。这成为开发可扩展视频编码最有价值的途径。

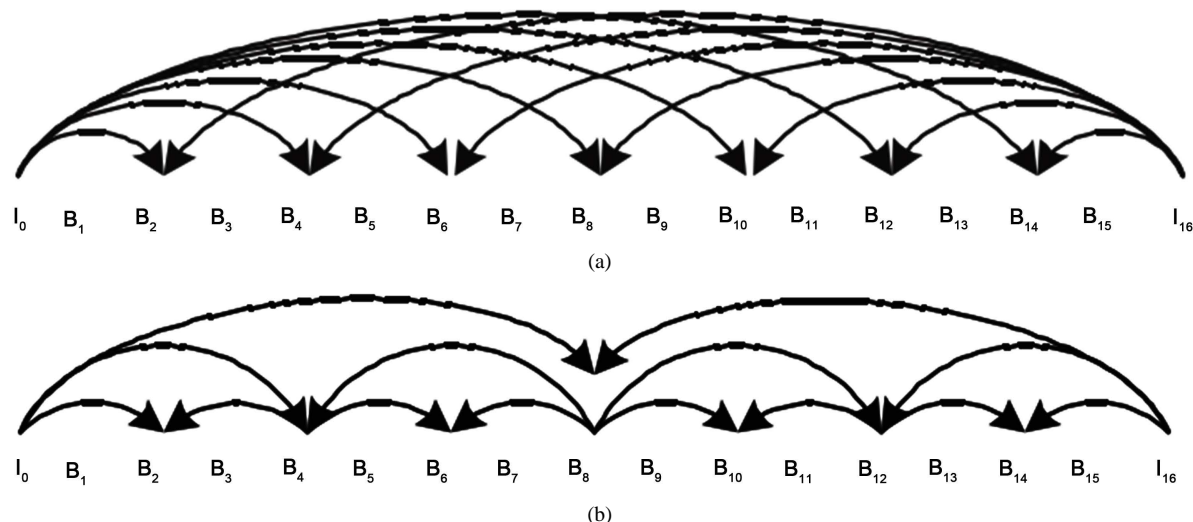


Figure 1. Illustration of classical and hierarchical B frame prediction structure. (a) Classical B frame prediction; (b) Hierarchical B frame prediction

图 1. 经典和分层 B 帧预测结构说明。(a) 经典 B 帧预测；(b) 分层 B 帧预测

H.264 SVC 在图 1(b)中揭示了层次 B 帧预测。图中显示通常采用二进位的层次 B 帧即在连续 I 或 P 帧之间有 $\beta = 2^\tau - 1$ ，(其中 τ 是一个正整数)。从图 1(b)可以看出帧 I0 和帧 B8 可以预测帧 B4。层次结构中，帧 B4 和帧 B8 可以预测帧 B6，反过来帧 B4 和帧 B6 可以预测帧 B5。

根据可预测的层次结构，利用时间的可伸缩性，B 帧被分配到 $\tau = \log_2(\beta + 1)$ 时间增强层。图 1(b)中，在连续的 I 帧之间有 $\beta = 15$ B 帧，I 帧属于时间基层 $T = 0$ ，帧 B8 属于第一个时间增强层 $T = 1$ ，帧 B4 和 B12 属于第二个时间增强层 $T = 2$ ，帧 B2, B6, B10 和 B14 属于第三层增强层 $T = 3$ ，剩下的 B 帧属于第四(最高)时间增强层 $T = \tau = 4$ 。因此，总共有 $\tau + 1$ 个时间层。

从图 1 所示，古经典的和层次 B 帧预测的编码顺序从根本上不同。一般来说，一个给定的帧 n 可以被编码之前，帧 n 所参考的所有帧都需要被编码，以便他们可以在编码形式上作为帧 n 编码参考。此帧编码顺序为 I0, I16, B1, B2, ...。这种编码限制了广义 B 帧概念。相比之下，分层 B 帧预测的帧编码顺序为 I0, I16, B8, B4, B2, B1, B3, ...。这些差异在帧编码顺序中意味着不同的网络传输延迟。

2.2. 可扩展层的 H.264 视频编码

H.264SVC 支持可扩展层编码，即提供时间可扩展性、空间可扩展性和质量(信噪比)可扩展性[3]。一个可扩展层的编码包含一个基本层和一个或多个增强层。H.264 SVC 最多可以支持 128 个层，而一个编码的实际层数取决于其应用。

1) 时间可扩展性：H.264 SVC 提供时间可扩展性，即适应性帧频，利用分层 B 帧结构。

2) 空间可扩展性：空间可扩展性提供了不同的空间帧决议，如，一个常见的帧间格式(CIF) 352×288 像素基础层和一个通过解码基础层和增强层得到的 704×576 满像素决议。每个空间层包含了运动补偿预测和内部预测，H.264 SVC 提高了编码效率。

3) 质量可扩展性：H.264 SVC 最多可提供 8 个质量层，可连续提高视频帧的信噪比。H.264 SVC CGS 采用相同的内部分层预测机制作为 H.264 的空间可扩展性编码，除了有关空间帧扩展决议的操作，

根据层间预测机制的概述，一个给定的 CGS 增强层依赖于低 CGS 增强层和基础层进行解码。此外，CGS 编码的每层都采用分层 B 帧预测结构。任何帧内编码的一个编码器与一个基本配置可以切换到解码更少或更多的 CGS 增强层。

4) 时间空间组合信噪比的可扩展性: H.264 SVC 标准支持组合可扩展性, 即视频流的提取与一个编码比特流中帧的不同组合频率、空间分辨率和信噪比的质量。

3. H.264 SVC 视频编码的结构

不同类型的 H.264 视频编码有不同的轨迹特征结构, 如基础的单层(非可扩展性)编码结构即视频编码可扩展层和可扩展子层等附加功能。通常有两种视频轨迹, 即视频帧尺寸和质量跟踪 1); 失真轨迹抵消 2)。

3.1. 单层 H.264 视频编码

通常, 视频跟踪库需要选择一个视频标题(序列)及编码的空间分辨率。1) 帧时间特征: Δ 表示帧周期, 则 $1/\Delta$ 表示频率。许多当前的视频帧速率是 30 帧/s (或 25 帧/s), 对应于一个帧周期的 $\Delta = 33.33$ ms (或 $\Delta = 40$ ms)。2) 帧大小表征: 用 X_q 表示, $n = 0, 1, \dots, N-1$, 表示帧大小(字节数)的视频帧 N 的编码与量化参数。3) 视频帧质量特性: 目前视频质量评价通过计算机信号处理技术实现。峰值信噪比(PSNR)能适度准确的评价视频质量。分别用 Q_q 、 Y_n , $n = 0, 1, \dots, N-1$, 表示质量峰值信噪比(PSNR)的亮度分量视频帧 N 编码与量化参数 q 。而视频跟踪可以提供额外的视频质量特征。4) 跟踪格式: 一个跟踪文件通常是一个文本文件的 ASCII 字符。包括标题部分和实际的视频跟踪表格的格式。标题给出一般的视频信息, 如视频标题、帧(空间)的分辨率和帧频, 以及它的编码, 量化参数等。表格部分为描述编码的视频帧, 包含一个编码帧的冗长。

3.2. H.264 SVC 视频编码的可扩展层

通常 H.264 视频编码可扩展层的每层轨迹与单层编码轨迹结构类似。每层的帧大小与编码层的字节大小对应。然而, 给定层轨迹的质量特征对应所有层的视频质量[4]。

时间可扩展性, 即帧频适应性, 可以灵活的与其它 H.264 SVC 可扩展性模式结合。帧大小和质量的痕迹为每组 $\tau + 1$ 各轨迹, 因此每个轨迹都有特性的帧特征。当频率对应基层 $T = 0$ 时, 给出 I 和 P 帧的大小, 其它所有帧的大小为 0。峰值信噪比 PSNR 值的偏移失真包括帧丢失在低时间层, 并在真内部成功解码的效果。

对应于时间层 $T = 1$ 的轨迹包含时间基层帧和 $T = 1$ 时间增强层, 有效地实现基层帧频的加倍。也就是说, $T = 0$ 和 $T = 1$ 时间层的帧的大小和峰值信噪比 PSNR 质量将高于高帧时间层 $T = 2, \dots, \tau$, 大小为 0 和质量为时间层 $T = 0$ 和 $T = 1$ 的失真抵消值。这种结构能应用到更高的时间层, 最高层 $T = \tau$ 包含所有帧大小和峰值信噪比 PSNR, 并对应于全帧视频编码序列的频率。

3.3. H.264 SVC 编码的子层可扩展性

优先级和目标比特率轨迹: 一个视频比特流与目标比特率是指从基层比特率到完整编码比特率可以从选定帧中提取一个 H.264 编码。视频数据流的提取与规定的目标比特率最大限度地利用视频质量的优化问题。如图 2 所示, 它使得基层帧依赖于不同帧的高层。

4. H.264 视频编码的流量、质量特征及评估网络传输

优先级水平评估与优化策略使用两路编码的比特流和原始(未编码)视频源。优化旨在最大化视频编码每帧的 PSNR, 同时满足给定目标比特率超过整个视频持续时间序列。优先级描述帧大小和 PSNR 帧品质对应于视频比特流规定的优先等级 P。对于给定的 H.264 SVC 编码索引的优先级别 P , $P = 63, \dots, 1, 0$ 。优先级 P 给每个帧 n 总大小和 PSNR 帧质量更高。

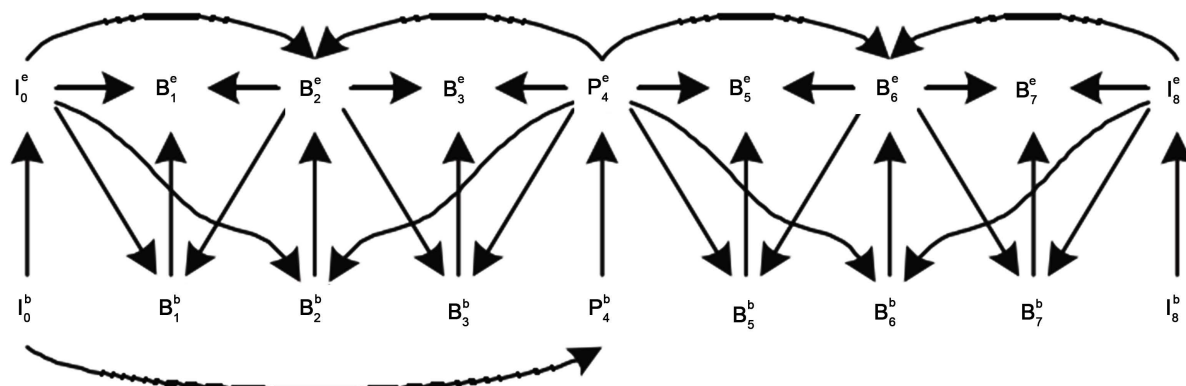


Figure 2. Sublayer-scalable H.264 SVC MGS coding

图 2. H.264 SVC 视频编解码子层扩展结构图

视频流量和质量特征的评价主要在于率失真(RD), 曲线和率可变性失真(VD)曲线。率失真(RD)曲线是视频质量 PSNR 平均亮度结点, 如式(1)所示:

$$\bar{Q}^{q,Y} = \frac{1}{N} \sum_{n=0}^{N-1} Q_n^{q,Y} \quad (1)$$

作为平均比特率函数, 这是通过平均帧尺寸公式得到, 如式(2)所示:

$$\bar{X}^q = \frac{1}{N} \sum_{n=0}^{N-1} X_n^q \quad (2)$$

帧周期为 \bar{X}^q/Δ , 视频跟踪中单个帧大小为 X_n^q 和质量 Q_n^q 通过量化参数 q 和率失真 RD 曲线中获取一系列量化参数。而平均亮度的峰值信噪比(PSNR)通过均方误差(MSE)值转换进行获得。

视频跟踪可以用来进行网络视频流量模拟, 如视频帧的播出时间延迟。帧的捕获时间很短, 每个帧的编码、传输和解码需要一个帧周期。如图 3 所示, 视频帧是由帧类型和瞬间捕获(记录)时间表示。I0 表示帧内编码(I)的捕获时间为 0, 而 B8 表示双向 B 帧编码捕获时间在 8Δ 。编码、传输和解码的时间在时间轴上分别得到显示。

如图 3(a)所示, 帧 B8 编码在帧周期中从 24Δ 到 25Δ , 传输帧周期从 25Δ 至 26Δ , 解码帧周期从 26Δ 到 27Δ 。同时可以看出视频编码、传输、捕获之间的时间偏移量。数据流请求有效时在 $\Delta 16$, 传输的编码帧 I0 则从 16Δ 开始, 视频开始播放时再推迟 3Δ 。从图 3(b)可以看出, 视频编码数据流与分层 B 帧之间相应时间是 6Δ , 即从 16Δ 到 22Δ 。而连续 I 和 P 帧之间的 βB 帧, 其在经典 B 帧时的视频数据流延迟为 $(4 + \beta)\Delta$, 而在分层 B 帧时的延迟为 $(3 + \beta = \log_2(1 + \beta))\Delta$ 。

H.264 SVC 编码在时间上具有可扩展性, 即可以动态调整帧频率, 通过丢弃和增加分层 B 帧的时间层 $T = 1, 2, \dots, \tau$ 。视频质量 PSNR 帧在层 $T = 0, 1, \dots, \tau - 1$, 可以读取帧大小和质量跟踪。在 $T = 0, 1, \dots, \tau - 1$ 层读取帧 PSNR 质量下降的失真偏移量。

目标比特率跟踪可以用来模拟传输视频数据流, 是符合规定的平均比特率的持续时间的完整视频序列编码[5]。而优先级 P 采用不同的方法优先传送视频包或者适应视频比特率。如高优先级视频数据与优先级级别范围在 63 到 P 可能会传播到更高优先级等级分化服务网络, 而视频数据与优先级级别 P1 与 0 是传播与更低的优先级。

基层帧 n 在网络传输的关键取决于其在帧预测结构中的位置。若帧 n 没有依赖帧, 其传输就不受依赖帧的影响; 若帧 n 依赖其他帧, 则帧 n 的质量 PSNR 受到所依赖帧的影响。

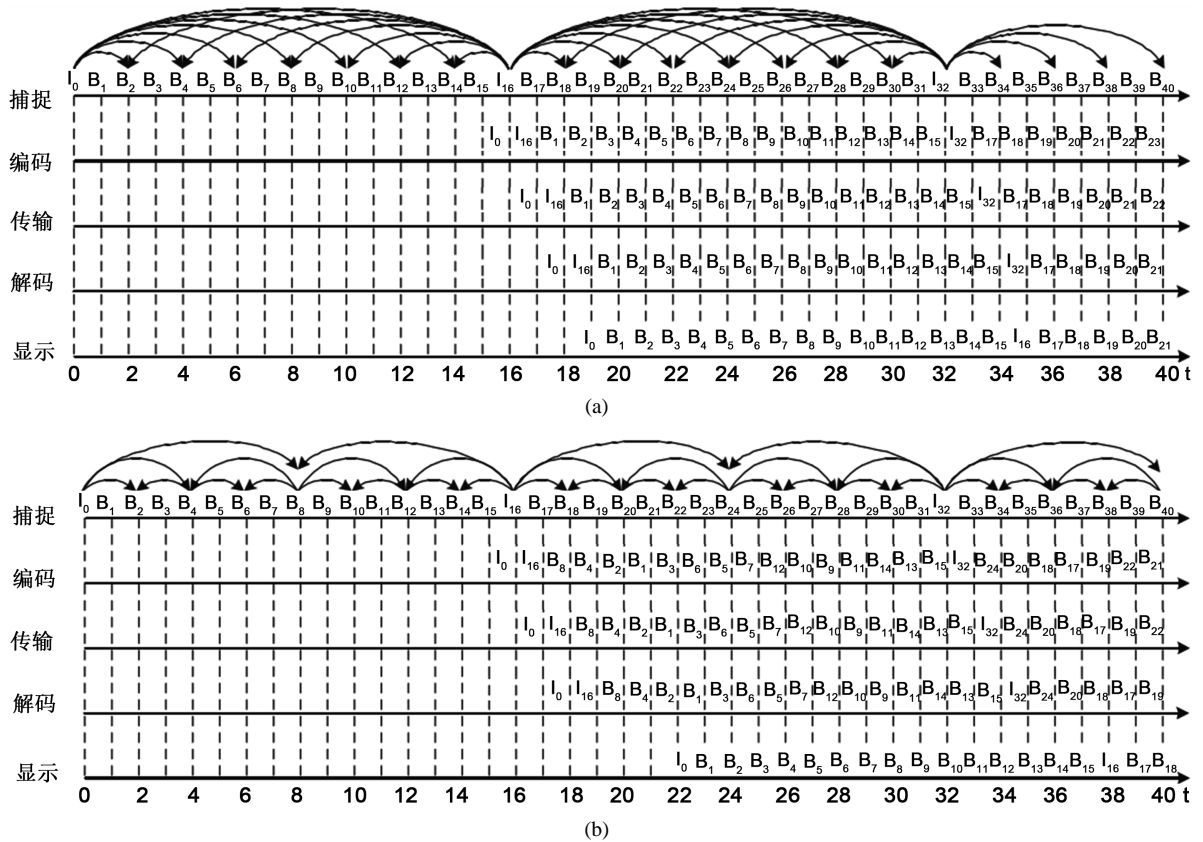


Figure 3. Delay comparison of classical and hierarchical B frame prediction. (a) Classical B frame prediction; (b) Hierarchical B frame prediction

图 3. 经典和分层 B 帧预测的时间延迟比较。(a) 经典 B 帧预测; (b) 分层 B 帧预测

5. 结论

论文全面阐述了 H.264 视频编码和 H.264 视频跟踪网络传输的评估方法。分别介绍了单层视频编码(非可扩展性), 及视频编码在时间、空间和质量(信噪比)、尺寸等方面的可扩展性。通过通信及网络等多方面知识理解 H.264 视频编码的视频网络传输。描述了视频编码跟踪结构, 帧大小(字节)和质量(PSNR)等不同类型的 H.264 编码视频, 介绍了视频数据流的生成, 并对 H.264 SVC 单层编码与层次 B 帧的率失真(RD)的性能进行了分析。而视频编码跟踪评估采用计算密集型的视频编码和解码, 限制了评估和长视频序列所需严格的统计结果。

参考文献 (References)

- [1] Wiegand, T., Sullivan, G.J., Bjontegaard, G. and Luthra, A. (2003) Overview of the H.264/AVC Video Coding Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, **13**, 560-576. <https://doi.org/10.1109/TCSVT.2003.815165>
- [2] Van der Auwera, G., David, P., Reisslein, M. and Karam, L.J. (2008) Traffic and Quality Characterization of the H.264/AVC Scalable Video Coding Extension. *Advances in Multimedia*, Article ID: 164027, 1-27. <https://doi.org/10.1155/2008/164027>
- [3] Schwarz, H., Marpe, D. and Wiegand, T. (2007) Overview of the Scalable Video Coding Extension of the H.264/AVC standard. *IEEE Transactions on Circuits and Systems for Video Technology*, **17**, 1103-1120. <https://doi.org/10.1109/TCSVT.2007.905532>
- [4] Amonou, I., Cammas, N., Kervadec, S. and Pateux, S. (2007) Optimized Ratedistortion Extraction with Quality Layers

in the Scalable Extension of H.264/AVC. *IEEE Transactions on Circuits and Systems for Video Technology*, **17**, 1186-1193. <https://doi.org/10.1109/TCSVT.2007.906870>

- [5] Li, X., Amon, P., Hutter, A. and Kaup, A. (2011) Performance Analysis of Interlayer Prediction in Scalable Video Coding Extension of H.264/AVC. *IEEE Xplore: IEEE Transactions on Broadcasting*, **57**, 66-74. <https://doi.org/10.1109/TBC.2010.2082370>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2163-3983, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: hjwc@hanspub.org