面向主观质量的视频传输编码参数的选择

杜 麟 田 畅 吴泽民 张兆丰 胡 磊 张 磊
(中国人民解放军理工大学通信工程学院 南京 210007)

摘 要 在视频传输中有两个重要的问题需要解决:1)在已知网络带宽以及丢包率的情况下如何分配信源和信道的编码速率;2)在信源编码速率限定的条件下如何选择编码的参数。针对问题 2),分析了现有的主观质量模型 VQMTQ,选择 VQMTQ 作为视频主观质量评价的指标。联合码率模型在给定目标码率时能够选择最佳的编码参数,使得编码后的视频主观质量尽可能高。该研究对于实际应用具有重要价值,保障了人眼的主观感受。

关键词 视频传输,编码参数,主观质量
中图法分类号 TN919.85 **文献标识码** A

Choice of Coding Parameters in Video Transmission for Perceptual Quality

DU Lin TIAN Chang WU Ze-min ZHANG Zhao-feng HU Lei ZHANG Lei
(College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China)

Abstract There are two important problems that need to be solved in video transmission. One is how to allocate the rate for source and channel coding under the known bandwidth and packet loss probability. Another one is how to choose the coding parameters under the restricted bitrate. This paper focused on the second problem. We analyzed the existed perceptual quality model VQMTQ, and chose VQMTQ as the index to evaluate the perceptual quality of videos. Combined with the rate model, the proposed method chooses the best coding parameters under the given target bitrate, which not only satisfy the limit of bitrate, but also make the compressed video have the best perceptual quality, which can be used in practice and protect the human perceptual feelings.

Keywords Video transmission, Coding parameters, Perceptual quality

1 引言

随着互联网技术的不断发展,视频应用在日常生活中越来越普及。用户可用户通过网页进行视频的点播服务(如优酷网、爱奇艺等),同时还可通过 QQ、微信等手机软件进行即时视频通话,因此研究视频传输保障具有重大意义。典型的视频传输系统包括 3 部分:视频的编码、网络中视频的传输以及视频的解码和播放。网络带宽的跨越式发展使得人们对于视频质量的要求越来越高,从最开始的标清到高清,再到现在的超清,意味着巨大的视频数据信息需要处理,因此在网络传输之前需要对视频进行压缩处理。HEVC[1-2] 是最新的视频编码标准,综合考虑了感知编码的特性,是目前研究的热点之一,但是 HEVC 的编码计算复杂性过高,难以实现实时处理,目前尚没有广泛的应用,因此本文仍采用 H.264 编码标准[3] 进行编码。H.264 视频编解码标准是由 ITU_T 视频编码专家组和 ISO/IEC 动态图像专家组组成的联合视频组提出的一种高度压缩的数字视频编解码标准。与此同时,网络传输往往在受限的带宽条件下进行,在进行视频压缩编码时通常需要考虑是采用原始的帧率不变但是以牺牲质量为代价来达到一定的压缩比,还是采用较小的帧率以保证较小的失真来实现压缩。因此视频编码参数的选择在视频编码中起着十分重要的作用。

视频编码的目的是在压缩视频数据量的同时保障解码后的视频质量,传统的视频传输保障使用端到端的失真来衡量整个视频传输系统,此类保障的目的是使得端到端的失真最小。因此,端到端失真的估计是丢包网络视频率失真优化以及信源信道联合编码的关键,准确的失真估计能够为视频编码以及网络传输过程中采用更为有效的编码和传输策略提供依据。端到端的失真分为信源编码失真与信道传输失真,文献[4-6]详细分析了视频传输系统中失真的估计;文献[7]结合一阶矩提出了一种新的端到端失真估计的递归计算模型;文献[8]重点研究了错误隐藏手段对视频传输的影响,采用了新的错误补偿手段使得端到端失真更为准确。在上述模型中失真的估计均基于 PSNR,PSNR 是常用的客观视频质量评价指标,其优点是计算简单,但是与人的主观感知相关性不够,因为人眼是最后的观测者,所以主观感知质量在视频编码和网络传输中起着至关重要的作用。近年来,随着主观感知质量的不断研究和发展,出现了一系列新的主观感知质量模型[9-10]。文献[11]用端到端主观感知质量的下降来衡量整个视频传输系统,将端到端的主观感知质量的下降分为信源端编码导致的主观感知质量下降 $PDMOS_C$ 和信道端传输导致的主观感知质量下降 $PDMOS_L$ 两部分,并分别研究了 PSNR 和丢包对 $PDMOS_C$ 和 $PDMOS_L$ 的影响,从而建立了相应的计算模型;文献[12]提出用视敏度表示视频的主观感知质量,

本文受国家自然科学基金(61501509)资助。

杜 麟(1992-),男,硕士生,主要研究方向为视频传输保障;田 畅(1963-),男,教授,主要研究方向为数据链系统与信息感知;吴泽民(1973-),男,副教授,主要研究方向为信息融合、图像识别;张兆丰(1991-),男,硕士生,主要研究方向为图像分析;胡 磊(1985-),男,讲师,主要研究方向为压缩感知、视频信息处理;张 磊(1972-),男,教授,主要研究方向为宽带网络与交换技术、嵌入式系统。

视敏度[13-14]是最基本的人眼视觉系统（Human Visual System，HVS）视觉处理机制，它表征了 HVS 对不同频率视觉信号的响应强度，主要有时域对比敏感度和空域对比敏感度；文献[15]主要研究了量化以及帧率对视频主观感知的影响，并提出了新的计算模型 VQMTQ；Wang 首先从时间域和质量域出发[16]，分别研究了帧率以及量化步长对主观感知质量和码率的影响，随后在文献[17]中加入分辨率对主观感知质量的影响，从而提出了联合的主观感知质量模型 Q-STAR。因此，本文提出了一种基于主观感知质量和码率模型的编码参数的选择方法，在码率受限的条件下选择最佳的编码参数从而保障编码后的视频质量。

本文第 2 节简单介绍了主观感知质量模型 VQMTQ；第 3 节对码率模型进行了简单的总结以及介绍了一种常见的码率模型；第 4 节详细介绍了基于主观感知质量和码率模型进行编码参数选择的方法，并通过实验进行验证；第 5 节总结全文并展望未来。

2　主观感知质量模型——VQMTQ

在文献[15]中，作者分别研究了视频帧率以及量化步长对视频主观感知质量的影响。采用不同的帧率以及量化步长对视频库中的序列进行压缩编码，对编码后的视频序列进行主观打分从而得到视频序列的 MOS 值，并进行分析。通过分析发现视频帧率和量化步长对视频主观感知质量的影响是相互独立的。考虑在实际编码过程往往采用自适应 QP 的方式，而 PSNR 能够准确地反映 QP 的变化，因此分别利用反指数函数和 Sigmoid 函数建立帧率和 PSNR 与主观感知质量之间的关系，从而提出了新的计算模型 VQMTQ，该模型的计算表达式如下：

$$TCF(f)=\frac{1-\mathrm{e}^{-b\frac{f}{f_{\max}}}}{1-\mathrm{e}^{-b}} \tag{1}$$

$$SQF(PSNR)=\hat{Q}_{\max}\left(1-\frac{1}{1+\mathrm{e}^{p(PSNR-PSNRt)}}\right) \tag{2}$$

$$VQMTQ(PSNR,f)=SQF(PSNR)TCF(f) \tag{3}$$

其中，$\hat{Q}_{\max}$ 表示同一视频序列采用不同编码参数进行编码后最高的主观感知质量，p 为与视频内容无关的参数；f 表示视频序列的帧率；$PSNR$ 表示编码后视频的 PSNR 值；$PSNRt$ 和 b 是与视频内容有关的模型参数，通过从视频序列本身提取特征来进行预测，特征提取如表 1 所列。

表 1

特征	特征描述
MDA	运动矢量的方向，$\theta_{MV}=\arctan(MV_X/MV_Y)$
DFD	基于运动估计的帧差
$\hat{G}_m$	Gabor 滤波的输出均值
FD	连续帧的帧差
STD	一帧图像像素值的标准差
NFD	$NFD=FD/STD$

模型参数的预测方程如下：

$$\hat{b}=10.72-0.6\cdot MDA-0.13\cdot DFD \tag{4}$$

$$PSNRt=34.8298-1.88\cdot\hat{G}_m-2.23\cdot NFD \tag{5}$$

3　码率模型

大多数对码率模型的研究都基于固定的帧率，如 MPEG-4 标准中提出的模型 $R=\frac{A_1}{q}+\frac{A_2}{q^2}$，其中 A_1 和 A_2 为模型参数；以及 H.264 参考软件中采用的码率控制机制 $R=MAD\cdot\frac{X_1}{QP}+X_2$，其中 MAD 表示信号预测之后的平均绝对误差，X_1 和 X_2 为模型参数。

文献[16]则从编码参数组合的角度出发，研究了视频帧率和量化步长对码率的影响。利用 H.264 标准参考软件 JM 对视频序列进行编码，此时主要调整视频的帧率以及编码的 QP 值。统计编码后视频序列的实际码率，通过分析发现，利用幂函数可以对码率值进行准确的建模，因此提出了一种码率计算模型，该模型的计算表达式如下：

$$R(q,f)=R_{\max}\left(\frac{q}{q_{\min}}\right)^{-a}\left(\frac{f}{f_{\max}}\right)^{b} \tag{6}$$

其中，$q_{\min}$，$f_{\max}$ 取决于实际编码中最小的量化步长以及视频的最大帧率；$R_{\max}$ 则为采用最小的量化步长以及最大的帧率进行编码时视频序列的实际码率；a 和 b 为模型参数，从视频序列本身提取特征值对 a，b 和 $R_{\max}$ 进行预测，特征值选取如表 2 所列。

表 2

特征	特征描述
μ_{FD}	相邻两帧之间帧差均值的平均
μ_{MVM}	运动矢量大小的均值
μ_{MDA}	θ_{MV} 的标准差的均值，$\theta_{MV}=\arctan(MV_X/MV_Y)$
$\eta(\mu_{MVM},\mu_{MDA})$	$\eta(\mu_{MVM},\mu_{MDA})=\mu_{MVM}/\mu_{MDA}$

模型参数的预测方程如下：

$$\begin{pmatrix} a \\ b \\ R_{\max} \end{pmatrix}=H\begin{pmatrix} 1 \\ \mu_{FD} \\ \mu_{MVM} \\ \eta(\mu_{MVM},\mu_{MDA}) \end{pmatrix} \tag{7}$$

其中，H 为预测系数矩阵。

4　编码参数的选择

在实际传输系统中，由于资源有限，需要考虑编码所产生的实际码率，同时还要兼顾视频的质量。当给定码率限制时，JM 采用一种自适应 QP 的方式进行编码，编码开始时采用最小的量化步长进行编码，当产生的码率超过限定值时即增大量化步长。这种方法的好处是可以提高视频的质量，但是在目标码率较低时，编码产生的实际码率往往大于目标码率。基于上述的主观感知质量 VQMTQ 模型以及码率模型，本文提出了一种视频编码参数的预选择方法。在给定目标码率时，选择最佳的编码参数组合使得编码后的视频主观感知质量尽可能高，问题的数学表示如下：

$$\max\ VQMTQ \tag{8}$$

$$\text{s.t.}\ R(q,f)\leqslant R_t \tag{9}$$

其中，R_t 为目标码率。

从式(1)和式(2)可以看出，VQMTQ 模型使用 PSNR 和帧率建立方程，能够计算编码后视频序列的主观感知质量，但在编码过程中由于 PSNR 无法计算，因此不能用于编码参数的选择，需要建立 PSNR 与编码参数之间的关系。本文主要考虑量化步长以及视频帧率的影响，需要建立 PSNR 与量化步长之间的关系。

利用 JM 软件采用不同的 QP 值对视频序列 CITY，CREW，HALL，HARBOUR，MOTHER 和 SOCCER 进行编码，并计算编码后视频序列的 PSNR 值，实验结果如图 1 所示。

由图 1 可知，对于 6 个视频序列，随着 QP 值增大，视频序列的 PSNR 逐渐变小，并且呈现出一种线性的关系，因此用线性函数 $y=k*x+b$ 对实验数据进行拟合。通过最小二乘

拟合得到的 k 值分别为 -0.01808，-0.01633，-0.0156，-0.02089，-0.0149 和 -0.01529，因此 PSNR 与 QP 之间的关系如下：

$$PSNR=k*(QP-28)+1 \tag{10}$$

其中，仅 k 为模型的参数，通过从视频序列本身提取特征进行预测。为了减小计算的复杂性，利用 VQMTQ 模型和码率模型中提取的特征值进行预测。为了用尽可能少的特征组合进行预测，采用交叉验证的方法验证各个特征值对模型参数的影响。从 6 个视频序列中任意选取 5 个用于训练，而剩下 1 个用于测试，计算预测结果与最小二乘拟合值之间的误差，通过误差值的大小进行特征组合的选取，实验结果如表 3 所列。

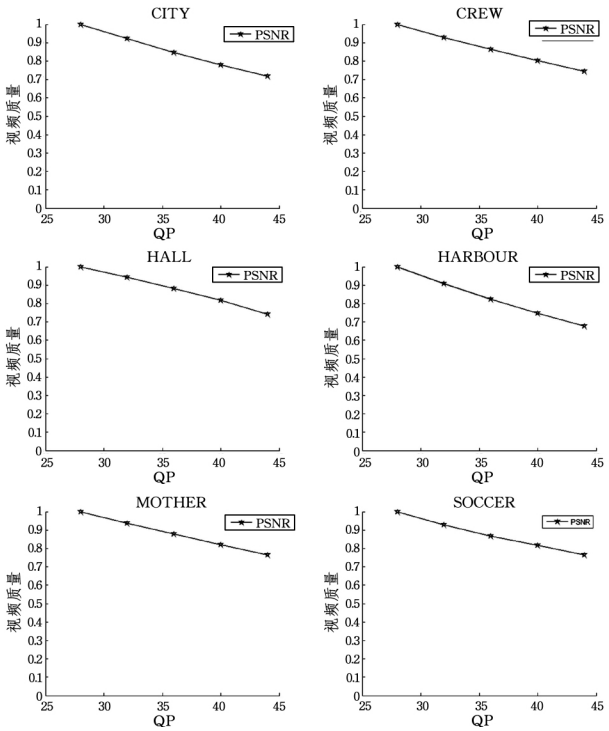

图 1 QP vs PSNR

表 3 特征组合的选取

	最佳组合	误差
$K=1$	G_m	0.0010
$K=2$	σ,G_m	0.0005
$K=3$	$\sigma,\eta(\mu_{FD},\sigma),G_m$	0.0008
K=4	$\mu_{FD},\sigma,\eta(\mu_{FD},\sigma),G_m$	0.0007

由表 3 可知，虽然当 $K=2$ 时预测的误差最小，但是采用不同个数的特征值进行预测所造成的误差相差不大，几乎可以忽略，为了用尽可能少的特征值进行预测，选取 $K=1$ 所对应的特征值。模型参数 k 的预测方程如下：

$$k=-0.0134-0.0010*G_m \tag{11}$$

将式(11)带入式(2)得到式(12)：

$$SQF(QP)=\hat{Q}_{\max}\left(1-\frac{1}{1+e^{p(k*(QP-28)+1-PSNRt)}}\right) \tag{12}$$

此时，主观感知质量 VQMTQ 模型和码率模型均是有关于 QP 和帧率的方程，所以能够进行编码参数的选择。

5 实验结果与分析

实验仿真平台为 Matlab 2010b，利用 Matlab 标准库中求解非线性多元函数最小值的函数 fmincon 进行编码参数的选择。测试视频序列为 CITY，CREW，HALL，HARBOUR，MOTHER 和 SOCCER。图 2 为 6 个视频序列的简图。

注：从左至右、从上至下依次为 CITY，CREW，HALL，HARBOUR，MOTHER 和 SOCCER。

图 2 6 个测试序列

实验中最小的 QP 值设为 28，且 QP 的取值范围为 $28\leqslant QP\leqslant 46$。原始视频序列的帧率为 30fps，通过下采样的方法分别得到帧率为 15，7.5 和 3.75 的视频序列。在目标码率 R_t 为 50k，100k，200k 和 400k 的情况下进行测试。将实验结果与H.264标准参考软件 JM 中的自适应码率控制进行比较，实验结果如图 3 所示。

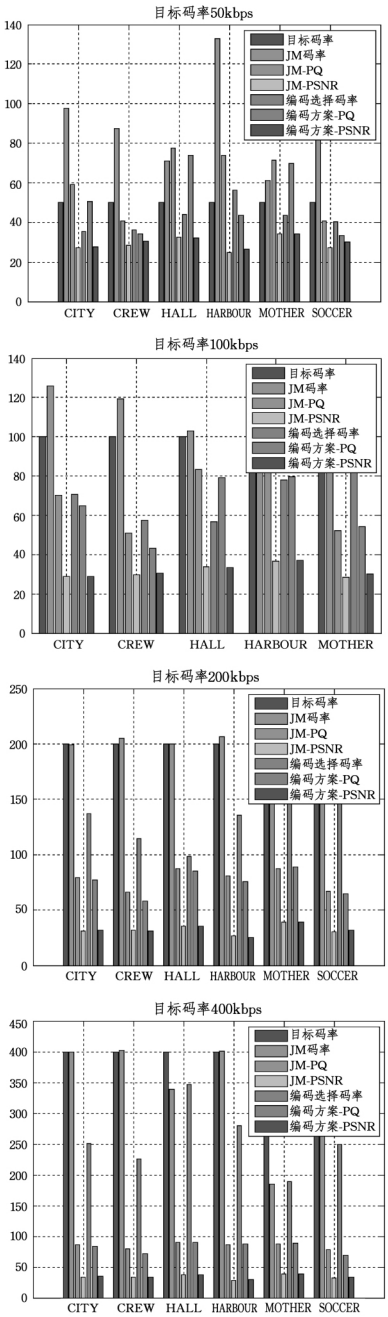

图 3 实验结果

图 3 中每一个视频序列对应的条状图从左至右依次为目标码率、JM 编码产生的实际码率、JM 编码后视频序列的感知质量、JM 编码后视频序列的 PSNR、本文提出的方法（预选择）编码产生的实际码率、预选择编码后视频序列的感知质量和与预选择编码后视频序列的 PSNR。

由图 3 的结果可知，在目标码率较低时（如 50K），采用 JM 中自适应码率控制的方法编码所产生的实际码率远远大于目标码率，特别是对于运动剧烈和场景复杂的视频序列。主要原因在于 JM 中自适应码率控制的方法采用先验知识预先设定帧率并保持帧率不变，仅仅通过改变每一帧编码的 QP 值来达到目标码率，而 QP 值的改变有限，当改变 QP 值仍不能达到目标码率时会出现跳帧的现象，从而导致在低目标码率时误差较大。而本文提出的基于主观感知质量 VQMTQ 模型和码率模型的编码参数预选择方案，无论是在低目标码率还是高目标码率时，基本上都能满足目标码率的限制。

另一个比较指标是视频的质量，分别计算了两类方法编码后视频序列的 VQMTQ 和 PSNR 值。从主观感知质量 VQMTQ 来看，按照本文提出的方法选择编码参数进行编码所得到的视频序列的 VQMTQ 与 JM 中自适应码率控制编码得到的视频序列的 VQMTQ 值相差不大。主要原因在于 JM 的码率控制保持原始的帧率不变，而为了满足目标码率的限制，本文提出的编码参数选择方法降低了视频序列的帧率，从而使得主观感知质量 VQMTQ 值偏低。从 PSNR 的角度来看，本文提出的编码参数选择方法使得编码后的视频质量在某些情况下甚至高于 JM。因此，从满足码率限制和主观感知质量尽可能高两方面考虑，本文提出的方法更优，能以更少的带宽资源实现相同的视频质量，具有实际使用价值。

结束语　网络视频应用越来越普及，保障用户接受的视频质量是最重要的任务，传统的基于 PSNR 的端到端最小失真保护从像素级出发，保障了传输过程中像素值的正确性，但是人眼的复杂决定了人眼视频感知的特征，使得传统的方法不能提供最佳的视觉效果。随着主观感知质量的不断研究，越来越多的主观感知质量模型被提出，因此本文提出了基于主观感知质量模型 VQMTQ 和码率模型的编码参数选择，使得编码后的视频序列具有最高的主观感知质量，并在 6 个序列上进行了测试，实验结果表明，本文提出的编码参数选择方案能有效地保障视频的主观感知质量。在实际应用中，对于离线视频点播服务，可以从整段视频序列中提取相应的特征进行模型参数的预测，从而选择最佳的编码参数进行编码；对于在线视频与监控视频的应用，可以从当前的 GOP 中提取视频特征进行模型参数的预测，从而选择最佳的参数作为下一个 GOP 的编码方案，每接收一个 GOP 即进行一次计算和更新。

但是本文所测试的视频序列格式均为 CIF，对于更高分辨率格式的视频序列未进行测试，这是接下来需要进行的工作。此外，本文提出的方法仅适用于视频编码端，能够有效保障编码后的视频主观感知质量。为了实现端到端的主观质量保障，接下来需要将主观感知质量模型用于传输过程，通过分析丢包和错误隐藏等对主观感知质量的影响，建立相应的主观感知质量模型。

参 考 文 献

[1] CHEN Z, LI Y. Recent advances in perceptual H. 265/HEVC video coding[C]// IEEE China Summit and International Conference on Signal and Information Processing. IEEE, 2015.

[2] ANEGEKUH L, SUN L, JAMMEH E, et al. Content-Based Video Quality Prediction for HEVC Encoded Videos Streamed Over Packet Networks[J]. IEEE Transactions on Multimedia, 2015, 17(8): 1.

[3] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, et al. Overview of the H. 264/AVC video coding standard[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2003, 13(7): 560-576.

[4] CHEN Z, WU D. Prediction of transmission distortion for wireless video communication: Algorithm and application[J]. Journal of Visual Communication & Image Representation, 2010, 21(8): 948-964.

[5] CHEN Z, WU D. Prediction of Transmission Distortion for Wireless Video Communication: Analysis[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2012, 21(3): 1123-1137.

[6] LI C, XIONG H, WU D. Delay-Rate-Distortion Optimized Rate Control for End-to-End Video Communication Over Wireless Channels[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2015, 25(10): 1.

[7] 张宏，景小荣，罗明敏. H. 264 视频传输中一种改进的端到端失真估计模型[J]. 重庆邮电大学学报（自然科学版），2014，26(3)：319-324.

[8] CHENG Q, AGRAFIOTIS D. End to end video distortion estimation with advanced error concealment considerations[C]// Visual Communications and Image Processing Conference. IEEE, 2014: 303-306.

[9] 蒋刚毅，朱亚培，郁梅，等. 基于感知的视频编码方法综述[J]. 电子与信息学报，2013，35(2)：474-483.

[10] LEE J S, EBRAHIMI T. Perceptual Video Compression: A Survey[J]. IEEE Journal of Selected Topics in Signal Processing, 2012, 6(6): 684-697.

[11] LIU T, WANG Y, BOYCE J M, et al. A Novel Video Quality Metric for Low Bit-Rate Video Considering Both Coding and Packet-Loss Artifacts[J]. IEEE Journal of Selected Topics in Signal Processing, 2009, 3(2): 280-293.

[12] BAIK E, PANDE A, STOVER C, et al. Video acuity assessment in mobile devices[C]// 2015 IEEE Conference on Computer Communications (INFOCOM). 2015: 1-9.

[13] LIN W, KUO C C J. Perceptual visual quality metrics: A survey [J]. Journal of Visual Communication & Image Representation, 2011, 22(4): 297-312.

[14] BALACHANDRAN A, SEKAR V, AKELLA A, et al. Developing a predictive model of quality of experience for internet video [J]. Acm Sigcomm Computer Communication Review, 2013, 43(4): 339-350.

[15] OU Y F, MA Z, LIU T, et al. Perceptual Quality Assessment of Video Considering Both Frame Rate and Quantization Artifacts [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2011, 21(3): 286-298.

[16] MA Z, XU M, OU Y F, et al. Modeling of Rate and Perceptual Quality of Compressed Video as Functions of Frame Rate and Quantization Stepsize and Its Applications[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2012, 22(5): 671-682.

[17] OU Y F, XUE Y, WANG Y. Q-STAR: a perceptual video quality model considering impact of spatial, temporal, and amplitude resolutions[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2014, 23(6): 2473-2486.