基于NetFlow的衡阳联通互联网流量的测量研究

王 琴[1],谭敏生[1,2]
(1.南华大学 计算机科学与技术学院,湖南 衡阳 421001;
2.南华大学 网络与信息安全研究室,湖南 衡阳 421001)

摘 要:从及时了解自身网络的负载状况和重要业务的带宽占用率,正确规划和评估网络的扩容、升级的角度出发,阐述了网络流量测量中的主要技术,重点介绍了NetFlow技术的相关原理。对衡阳联通互联网的拓扑结构进行了详细的分析,提出了基于NetFlow的衡阳联通网络流量测量方案,并分析了测量结果。实验结果表明,通过NetFlow流量测量,可以方便、有效地进行网络管理,为网络管理员维护网络提供了良好的工具。

关键词:网络流量;测量方案;NetFlow
中图分类号:TP393 **文献标识码**:A **文章编号**:1673-629X(2009)05-0122-05

Research on Network Traffic Measurement of Hengyang Unicom Based on NetFlow

WANG Qin[1], TAN Min-sheng[1,2]
(1. School of Computer Science and Technology, University of South China, Hengyang 421001, China;
2. Network and Information Security Lab, University of South China, Hengyang 421001, China)

Abstract: Mainly know the burden conditions of network and bandwidth occupancy rates of important businesses immediately, so can program and evaluate the construction and updating of network correctly. Introduces the main techniques about network traffic measurement, especially about the theories of NetFlow. Analyzes the structure of Hengyang Unicom network detailedly, and offers the network traffic measurement solution of Hengyang Unicom based on NetFlow, then analyzes the measuring results. The result indicates that network manager can do the network expediently and effectively according to network traffic measurement solution based on NetFlow.

Key words: network traffic; measurement solution; NetFlow

0 引 言

随着宽带互联网在中国的迅速发展,网络规模不断扩大,网络结构日渐复杂。为了更好地服务客户,通信运营商需要及时了解自身网络的负载状况和重要业务的带宽占用率,从而才能对网络所承载的各类业务进行及时、准确的分析以及正确规划和评估网络的扩容、升级。网络流量数据为网络的运行和维护提供了重要信息,这些数据对网络的资源分布、结构规划、服务质量、安全管理都比较重要。因此,对互联网的流量测量和分析是研究互联网的有效方法。

以衡阳联通互联网为平台,阐述了Cisco的网络流量测量技术NetFlow,从数据采集策略、端口选择、采样配置和设备配置等方面设计了网络流量测量方案。将该方案应用到联通互联网上,并以此为基础对采集的数据进行统计分析,为网管人员在网络管理中进行决策提供了有力的依据。

1 网络流量测量技术

网络流量测量主要是从网络设备上采集数据,获得流量数据后对其进行统计,并且存储网络流量的主要历史数据,定期形成性能报表[1]。网络管理员根据报表上的数据就可对网络的主要性能进行分析,通过分析性能的变化趋势以及网络性能的瓶颈问题,更好地进行路由和负载的设计。通过对网络流量的测量,可以决定网络拥塞控制,进而降低因网络拥塞带来的信息丢失和延迟,充分利用网络资源,提高服务质量。根据测量工具是否向被测量网络中发送数据进行划分,分为主动测量和被动测量[2,3]。

收稿日期:2008-09-09
基金项目:湖南省科技计划项目(2006GK3084);湖南省教育科学研究项目(05C487);衡阳市科技计划项目(2005KG01-015)
作者简介:王 琴(1983-),女,湖南衡阳人,硕士研究生,研究方向为计算机网络与信息安全;谭敏生,教授,硕士,硕士生导师,研究方向为计算机网络与信息安全。

主动测量是在选定的测量点上利用测量工具有目的地主动注入网络,并根据测量数据流的传送情况来分析网络的性能[4]。这意味着主动测量过程中会产生新的网络流量。主动测量具有响应速度快、适应性强的优点,对测量过程的可控性比较高,灵活、机动,易于进行端到端的性能测量。同时主动测量技术还可以探测网络的特定现象,如发现许多 Internet 端至端的延迟分布具有重尾特征。主动测量的缺点是给网络增加了潜在的负担,如果该测量未经仔细设计,可能会引起"Heisenberg(蝴蝶效应)",即额外的流量可能会干扰网络,对网络造成较大的影响,并使结果分析产生偏差。例如,在一个网络中,测量瓶颈链路带宽时采用的是发送大量大小不同的包获得延迟差的方法,那么随之产生的额外流量可能会阻塞网络路径。在测量包多数要排队的情况下,这种方法测出的瓶颈链路带宽是不准确的。

与主动测量相对应的是被动测量。被动测量是大多数测量工具采用的方法,在测量过程中,测量工具安装在网络中的某一个点上收集流量信息,如使用路由器或交换机收集数据,或者一个独立的设备被动地监测网络链路的流量,然后使用包过滤器捕获通过该点的数据包。因为包过滤器能够捕获网络流量而不会对网络造成什么影响(如果它们可以用本地磁盘记录流量),所以使用被动测量可以消除额外的流量负载和 Heisenberg 效应,对网络的行为没有影响[5]。这具有一定的优势,使得人们更倾向于使用被动测量。被动测量的缺点在于其基本上是基于对单个设备的监测,很难对网络端到端的性能进行分析,另外还存在用户数据泄漏等安全和隐私问题。但是被动测量较适合用来进行流量测量。

2　NetFlow 概述

目前比较常见的被动测量技术有基于侦听网络数据包的测量技术、基于 SNMP 的路由器流量测量技术、基于网络探针的流量测量技术和基于 flow 的流量测量技术[6]。NetFlow 是 Cisco 公司在 1996 年开发出来的技术,它既是一种交换技术,也是一种流量分析技术。目前主要应用的 NetFlow 技术版本为 NetFlow Version 5,该版本采集到的流量数据较详细,可以支持不同维度的统计分析。因此,文中采用的是 NetFlow 协议的版本 5。

NetFlow 是 Cisco 公司开发的用于进行流量测量的技术。NetFlow 的核心是利用了流(Flow)的概念,经常被用于以下几个方面:流量分析和监控;根据流量进行计费;实现网络加速;用于网络安全性分析。为了网络测量的方便,Cisco 的路由器中包含了流量监测模块,称为 NetFlow services[7]。开启 NetFlow 后,路由器在转发分组的同时,会根据配置,记录经过此路由器的分组情况,记录的内容可以包括源 IP 地址和目的 IP 地址、下一跳地址、输入和输出物理端口号、某个流的包数、某个流的总字节数、流的起始时间和结束时间、源和目的的协议端口号、协议类型、服务类型(TOS)。同时,每经过指定的时间,就把这些数据以一定的格式发送给指定的主机进行处理。依据网络的部署情况进行适当配置,将发送流量记录的目的地址设置为采集和分析器的地址,并指定端口。在指定的主机(充当采集和分析器)上安装采集和分析软件监听此端口。当有数据发送过来时,根据规定的数据包的格式对该数据包进行分析、记录和处理,这样就可以得到经过这些路由器的网络流量情况。NetFlow 交换技术在网络设备的网络层实现高性能的交换,它提供了一个高效的机制,可以用来处理安全访问列表,从而不必像其他交换方式那样,为完成同样的任务而付出较高的性能代价。

NetFlow 技术的实现至少需要 3 个模块,即源设备、采集器、分析器,工作架构如图 1 所示。其中源设备送出数据,采集器采集数据并存储到服务器上,以便数据分析器进行处理。

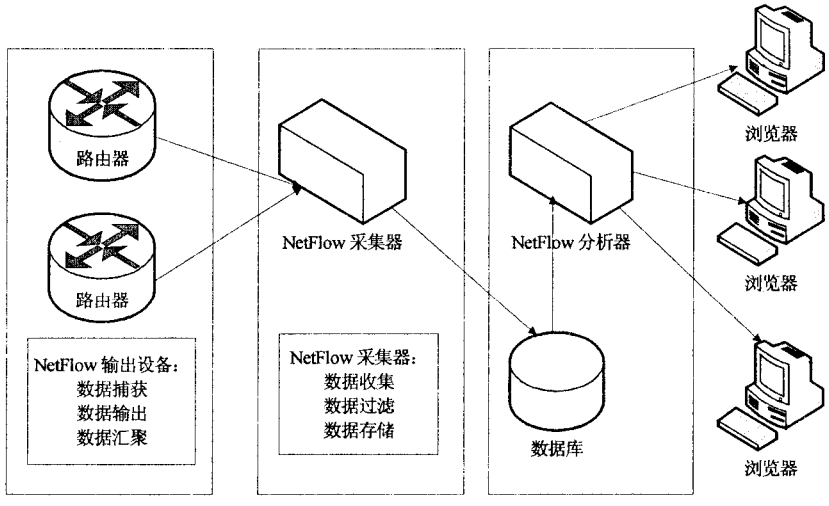


图 1　NetFlow 工作架构图

3　衡阳联通互联网流量的测量方案

3.1　衡阳联通互联网的拓扑结构分析

优良的拓扑结构是网络稳定运行的基础,要对网络流量进行分析,必须首先对网络的拓扑结构进行分

析。以衡阳联通互联网为例，来说明网络流量分析的方法。衡阳联通互联网的拓扑结构如图 2 所示。

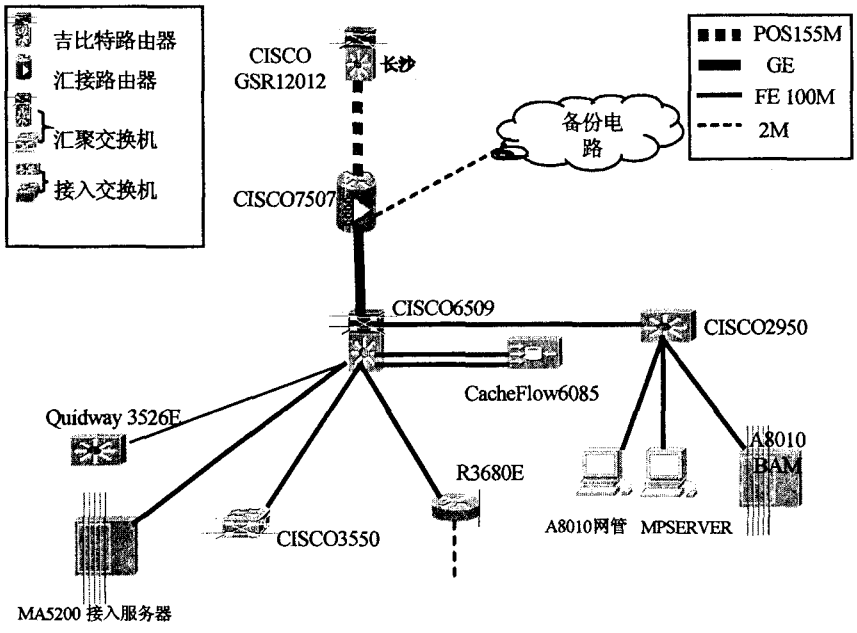


图 2 衡阳联通互联网拓扑结构图

根据衡阳联通互联网的实际情况，将网络拓扑结构分成了三个层次，如图 3 所示：核心层、汇聚层和接入层。核心层提供数据的高速交换；汇聚层是数据包的逻辑交换；接入层是终端用户的汇聚点[8]。

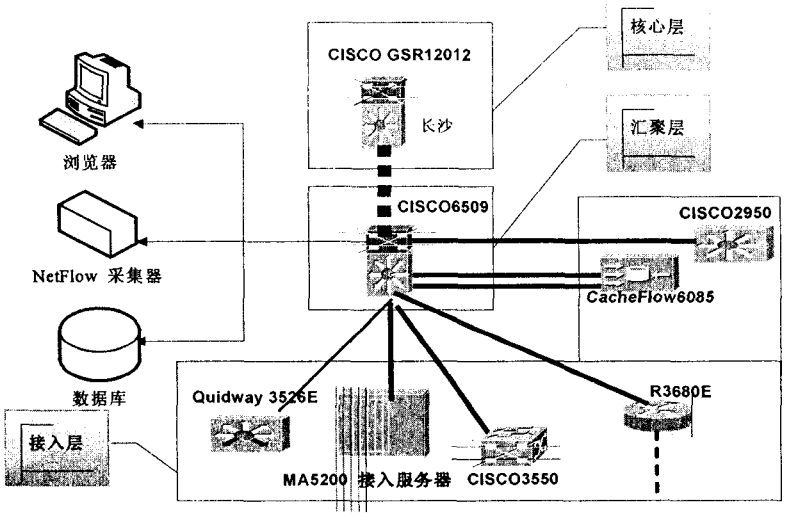


图 3 流量测量分层结构图

3.1.1 核心层

核心层是把各个汇聚层连接起来，为汇聚层提供高速率的数据交换，同时实现与各运营商骨干网络的互联，提供城市的高速数据出口。核心层的网络拓扑架构必须重点考虑其可靠性、可扩展性和开放性，所配置的设备要求其具有强路由功能。衡阳联通互联网的核心层是在长沙网络中心，由长沙 CISCO12012 设备实现，包括中心数据库和 web 服务器，负责网络流量数据的后期处理、整合、存储和 web 端发布。

3.1.2 汇聚层

汇聚层控制路由表的大小，收敛网络数据流量，将大量低速的链接接入核心层，以实现通信量的收敛，使核心层与汇聚层的连接最小化，同时减少核心层设备的路由路径的数量。在主要路由器和关键设备上设置采集分析器，负责对相应管辖区内网络设备的定期采集，收集、加工、处理相应数据的信息并将有价值的信息上传到网络中心。衡阳联通互联网的汇聚层是由衡阳网络中心的 CISCO6509 实现。

3.1.3 接入层

接入层通过线路资源以及接入设备将用户连入交换机，并提供多种业务的用户接入，比如专线用户、宽带用户、VPDN 用户等。接入层要确保用户的网络流量汇入网络，并且接入层路由器所接收的链接数不能超出其所承受的链接数。由于接入层是用户接入网络的入口，所以也是黑客入侵的首选目标。

3.2 衡阳联通互联网的流量测量方案

3.2.1 基于网络数据流(NetFlow)的端到端数据采集

从衡阳联通互联网的拓扑图中可以看出，数据网中所有数据都必须经过 Cisco CISCO6509 交换机，因此，在 CISCO6509 交换机所有使用业务的端口上打开了 NetFlow 协议。

如图 3 所示，流量测量系统配置结构采用集中采集的部署方式。当前衡阳联通互联网规模不是很大，将中央数据库、流量监控服务和流量采集都放置在衡阳网络中心，这样就能满足现在网络管理的需要，运行也比较稳定。考虑到随着衡阳联通的业务不断扩大，要监控的网络规模也会不断增大，当网络规模增大到一定程度时，网络中心将不能负荷中央数据库、流量监控和流量采集这么多的功能，将根据流量情况在接入节点分别进行流量的数据采集，这样可以避免对网络中心骨干链路造成压力。

3.2.2 NetFlow 的配置

NetFlow 的数据输出要求先在路由器或交换机上

定制 NetFlow 流输出,并选择输出流的版本、个数、缓存区的大小等,配置相应 NetFlow 流收集器的 IP 地址、端口等信息。另外,需要在 NetFlow 收集器端配置接收端口号、设置汇聚、过滤策略、格式等。NetFlow 的配置过程如下[7]:

router # enable
Password: * * * * *
Router # configure terminal //进入全局配置模式
Router (config) # interface FastEthernet 0/1
Router (config - if) # ip route - cache flow //指定流交换
Router (config) # exit
Router (config) # ip flow - export destination 192.168. * . * 9996 //将流量数据导入到目标机器 192.168. * . * 的 9996 端口上
Router (config) # ip flow - sampling - mode packet - interval 100 //该实例中采用 sampled 模式,采样间隔为 100:1.
Router (config) # ip flow - export source FastEthernet 0/1 //指定输出 NetFlow 数据源 f0/1 端口
Router (config) # ip flow - export version 5 //指定 NetFlow 的版本号 V5
Router (config) # ip flow - cache active timeout 30 //改变激活的超时时间,该值可以是 1 到 60 秒.默认为 30 分钟
router # write

配置完成后,可以使用如下命令查看 NetFlow 的输出情况:

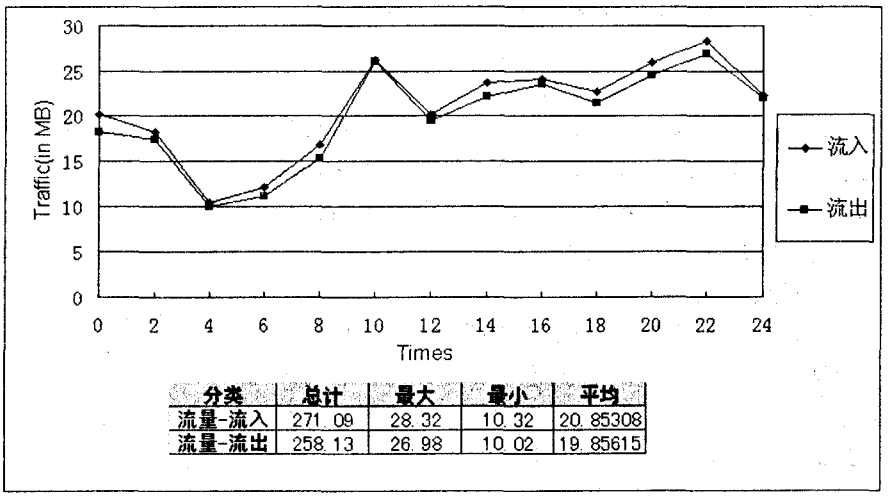
router # show ip flow export //查看 NetFlow 的输出信息
router # show ip cache flow //查看 NetFlow 统计信息

这样设置后,安装在 192.168. * . * 的 NetFlow 分析器就可以进行流量分析了。

4　测量结果与分析

4.1　同一天内不同时间段的网络流量测量与分析

2008 年 3 月 30 日上午 8 点,在 CISCO6509 上开启 NetFlow,进行实时的流量测量。图 4 所示为 2008 年 4 月 1 日一天的网络流量变化。网络流量的高峰期为 8:00—10:00 和 20:00—22:00,网络流量的低峰期为 3:00—5:00。因此在每天的 8:00—10:00 和 20:00—22:00(人们上网高峰期)可以采用增加路由的方式来缓解资源的紧张,在闲时为避免浪费资源,可以采用减少相应路由的方式来达到资源的充分利用。

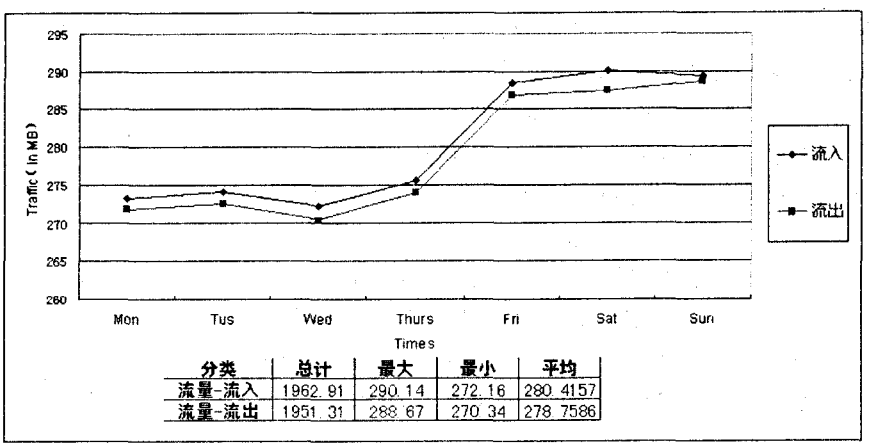


分类	总计	最大	最小	平均
流量-流入	271.09	28.32	10.32	20.85308
流量-流出	258.13	26.98	10.02	19.85615

图 4　一天的网络流量图

4.2　一周内的网络流量测量与分析

如图 5 所示,呈现了 CISCO6509 从 2008 年 3 月 31 日(星期一)至 4 月 6 日(星期天)一周的网络流量变化。周末的网络流量总是比较大,在周六(4 月 6 日)达到最大,明显高于其他时段的流量。网络管理员实时监控网络流量,当晚忙时的链路负荷达到门限值时,可采取增加电路的方式来缓解高峰期所带来的资源紧张问题。如果发现网络出现了异常大的流量,表明网络有可能受到攻击。可通过 NetFlow 采集器查看源 IP 地址与目的 IP 地址,及时采取相应的解决方案[9]。目前的解决方案有以下几种:



分类	总计	最大	最小	平均
流量-流入	1962.91	290.14	272.16	280.4157
流量-流出	1951.31	288.67	270.34	278.7586

图 5　一周的网络流量图

(1)切断连接。

如果能够确定异常流量的源地址以及源地址设备处于可控的情况下,最直接的解决办法是切断异常流量源设备的物理连接。

(2)过滤。

采用 ACL(Access Control List)过滤能够灵活实现针对源 IP 地址、目的 IP 地址、协议类型、端口号等各

种形式的过滤,但同时也存在消耗网络设备资源的副作用,下面为利用 ACL 过滤 UDP1434 端口的实例:

```
Access-list 101 deny udp any any eq 1434
Access-list 101 Permit ip any any
```

此过滤针对蠕虫王病毒(SQL Slammer),但同时也过滤了针对 SQL Server 的正常访问,如果要保证对 SQL Server 的正常访问,还可以根据病毒流数据包的大小特征实施更细化的过滤策略。

(3)静态空路由过滤。

能确定异常流量目标地址的情况下,可以用静态路由把异常流量的目标地址指向空(Null),这种过滤几乎不消耗路由器系统资源,但同时也过滤了对目标地址的正常访问,配置实例如下:

```
Ip route 192.168.*.* 255.255.255.255 Null
```

对于多路由器的网络,还需增加相关动态路由配置,保证过滤在全网生效。

5 结束语

实时了解网络的运行状况对通信运营商来说是必要的。合理地利用资源不仅可以减轻网络的维护难度,而且可以降低成本。对 NetFlow 测量技术进行了详细的介绍,并在衡阳联通互联网上部署了 NetFlow,对联通互联网进行监测。研究表明,通过流量测量,可以方便、有效地进行网络管理,为网络管理员维护网络提供了良好的工具。

参考文献:

[1] Williamson C. Internet Traffic Measurement[J]. IEEE Internet Computing,2001(10/11):70-74.

[2] 程 光,龚 俭.大规模高速网络流量测量研究[J].计算机工程与应用,2002,38(5):17-22.

[3] I-eland W E, Wilson D V. High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection[C]// in Proc. JEEEINFOCOM'91. Bal Harbour, FL:[s.n.],1991:1360-1366.

[4] 张 峰,雷振明.高速网络流测量及模型研究[J].计算机工程与应用, 2004,40(17):28-30.

[5] 王红莲.因特网网络流量测量技术的研究和实现[D].北京:北京邮电大学,2004.

[6] 潘 飞,高 岭.网络测量及其关键技术[J].计算机技术与发展, 2006,16(7):99-101.

[7] Cisco Systems. NetFlow service and Application White Paper [EB/OL].2001-06.http://www.cisco.com.

[8] 顾 静.电信数据网流量监测系统的设计与实现[D]. 长春:吉林大学,2007.

[9] 赵晓峰,徐义东.基于 NETFLOW 与 SNMP 的园区网流量监控系统[J].计算机技术与发展, 2008,18(5):168-171.

(上接第 121 页)

4 结束语

在图像的子带分解过程中,Á Trous 小波变换的移不变性解决了以往 Mallat 小波变换时由于抽取导致的相位失真的问题;Curvelet 变换的"各向异性"能更有效地保留原高分辨率图像的边缘特征;针对低频和高频部分采取的不同的融合规则,避免了小目标色彩和细节信息的丢失,有利于细节信息与背景色的分离。

文中的方法对空间分辨率之比在 1:4 以内的多光谱图像和全色图像进行融合时能得到较好的融合结果,在以后的研究中,需要适当选择合适的融合规则来融合空间分辨率之比较大的多光谱图像和全色图像。

参考文献:

[1] 梅安新,彭望琭,秦其明,等.遥感导论[M].北京:高等教育出版社,2001.

[2] POHL C, Van Genderen J L. Multisensor image fusion in remote sensing: concepts, methods and applications[J]. International J. Remote Sensing,1998,19(5):823-854.

[3] Nencini F, Garzelli S, Baronti S. Remote sensing image fusion using the curvelet transform[J]. Information Fusion, 2007,8(2):143-156.

[4] 王振飞,施保昌,王能超. 基于曲波变换的图像融合方法[J]. 小型微型计算机系统,2007,28(3):533-535.

[5] 田 闯,刘文波.基于 Curvelet 多聚焦图像融合[J].计算机技术与发展,2008,18(7):29-31.

[6] 那 彦,焦李成.基于多分辨率分析理论的图像融合方法[M].西安:西安电子科技大学出版社,2007.

[7] Candes E J. Ridgelets: theory and application[D]. USA: Standford University, 1998.

[8] Candes E J. Monoscale ridgelets for the representation of images with edges[R]. USA:Standford University,1998.

[9] 刘贵喜,陈文锦,杨万海. 融合参数对对比度塔形分解图像融合方法性能的影响研究[J]. 电力与系统学报,2006, 11(1):41-42.

[10] Shi WenZhong, Zhu ChangQing, Tian Yan, et al. Wavelet-based image fusion and Quality assessment[J]. International Journal of Applied Earth Observation and Geoinformation, 2005,6(3):241-243.

[11] Alparone L, Baronti S, Garzelli A. A global quality measurement of pan-sharpened mulitspectral imagery[J]. IEEE Geoscience and Remote Sensing Lett., 2004, 1(4): 313-317.