

Social Street View: Blending Immersive Street Views with Geo-tagged Social Media

Ruofei Du and Amitabh Varshney

The Augmentarium | UMIACS | University of Maryland, College Park

{ruofei, varshney} @ cs.umd.edu

www.SocialStreetView.com

Introduction

Social Media



image courtesy: plannedparenthood.org



Introduction
Social Media - Versatility





EXIF

Introduction

Meta data - Time of Creation



image courtesy: Brian Clegg

Introduction

Meta data - Location of Creation



image courtesy: squarespace.com

Introduction

Meta data - Camera Parameters

Related Work

Linear narrative visualization



image courtesy:
instagram.com,
facebook.com,
twitter.com

Related Work

Natural Immersive Virtual Reality?



image courtesy: drivenhealthy.com

Related Work

Karnath et al. and Loomis et al.

Related Work

Karnath et al. and Loomis et al.

Spatial awareness is a function of the temporal not the posterior parietal lobe

Hans-Otto Karnath, Susanne Ferber & Marc Himmelbach

Department of Cognitive Neurology, University of Tübingen, Hoppe-Seyley-Strasse 3, 72076 Tübingen, Germany

Our current understanding of spatial behaviour and parietal lobe function is largely based on the belief that spatial neglect in humans (a lack of awareness of space on the side of the body contralateral to a brain injury) is typically associated with lesions of the posterior parietal lobe. However, in monkeys, this disorder is observed after lesions of the superior temporal cortex¹, a puzzling discrepancy between the species. Here we show that, contrary to the widely accepted view, the superior temporal cortex is the neural substrate of spatial neglect in humans, as it is in monkeys. Unlike the monkey brain, spatial awareness in humans is a function largely confined to the right superior temporal cortex, a location topographically reminiscent of that for language on the left². Hence, the decisive phylogenetic transition from monkey to human brain seems to be a restriction of a formerly bilateral function to the right side, rather than a shift from the temporal to the parietal lobe. One may speculate that this lateralization of spatial awareness parallels the emergence of an elaborate representation for language on the left side.

Spatial neglect is a characteristic failure to explore the side of space contralateral to a brain lesion. Patients with this disorder behave as if one side of the surrounding space had ceased to exist. Since the early post-mortem studies, we have believed that, in humans, lesions located predominantly in the posterior parietal lobe are critical for this disorder. Analyses of computerized tomography scans of right-hemispheric stroke patients with neglect found that superimposed lateral projections of these lesions centred on the inferior parietal lobule (IPL)^{3,4} and the temporo-parieto-occipital (TPO) junction⁵. More recent studies have confirmed the validity of this conclusion although evidence for additional pathology leading

Related Work

Karnath et al. and Loomis et al.

Spatial awareness is a function of the temporal not the posterior parietal lobe

Hans-Otto Karnath, Susanne Ferber & Marc Himmelbach

Department of Cognitive Neurology, University of Tübingen, Hoppe-Seyley-Strasse 3, 72076 Tübingen, Germany

Our current understanding of spatial behaviour and parietal lobe function is largely based on the belief that spatial neglect in humans (a lack of awareness of space on the side of the body contralateral to a brain injury) is typically associated with lesions of the posterior parietal lobe. However, in monkeys, this disorder is observed after lesions of the superior temporal cortex¹, a puzzling discrepancy between the species. Here we show that, contrary to the widely accepted view, the superior temporal cortex is the neural substrate of spatial neglect in humans, as it is in monkeys. Unlike the monkey brain, spatial awareness in humans is a function largely confined to the right superior temporal cortex, a location topographically reminiscent of that for language on the left². Hence, the decisive phylogenetic transition from monkey to human brain seems to be a restriction of a formerly bilateral function to the right side, rather than a shift from the temporal to the parietal lobe. One may speculate that this lateralization of spatial awareness parallels the emergence of an elaborate representation for language on the left side.

Spatial neglect is a characteristic failure to explore the side space contralateral to a brain lesion. Patients with this disorder behave as if one side of the surrounding space had ceased to exist. Since the early post-mortem studies, we have believed that humans, lesions located predominantly in the posterior parietal lobe are critical for this disorder. Analyses of computerized topography scans of right-hemispheric stroke patients with neglect that superimposed lateral projections of these lesions centred on inferior parietal lobule (IPL)^{3,4} and the temporo-parieto-occipital (TPO) junction⁵. More recent studies have confirmed the validity of this conclusion although evidence for additional pathology

Immersive virtual environment technology as a basic research tool in psychology

JACK M. LOOMIS and JAMES J. BLASCOVICH
University of California, Santa Barbara, California

and

ANDREW C. BEALL
Massachusetts Institute of Technology, Cambridge, Massachusetts

Immersive virtual environment (IVE) technology has great promise as a tool for basic experimental research in psychology. IVE technology gives participants the experience of being surrounded by the computer-synthesized environment. We begin with a discussion of the various devices needed to implement immersive virtual environments, including object manipulation and social interaction. We review the benefits and drawbacks associated with virtual environment technology, in comparison with more conventional ways of doing basic experimental research. We then consider a variety of examples of research using IVE technology in the areas of perception, spatial cognition, and social interaction.

Human history records a progression of artifacts for representing and recreating aspects of external reality, ranging from language, drawings, and sculpture in earlier times to the more modern artifacts of photographs, movies, television, and audio recordings. Relatively recently, the digital computer and its associated technologies, including three-dimensional (3-D) graphics, have given rise to increasingly realistic artifacts that blur the distinction between reality and its representation (Ellis, 1995).

The ultimate representational system would allow the observer to interact "naturally" with objects and other individuals within a simulated environment or "world," an experience indistinguishable from "normal reality." Although such a representational system might conceivably use direct brain stimulation in the future, it will more likely use digitally controlled displays that stimulate the human sensory organs, the natural conduits to the brain. Displays of this type, referred to as *virtual displays* (VDs), although far from ideal, exist today. Following the terminology of others (e.g., Durlach & Mavor, 1995; Stanvick & Salvendy, 1998), we refer to the corresponding environment represented and stored in the computer and experienced by the user as a *virtual environment* (VE). *Virtual environment technology* (VET) refers inclusively both to VDs and to the VEs so created, including VEs produced by using conventional desktop computer displays.

(*Virtual reality* is widely used as an alternative term, but we prefer VE.) An immersive virtual environment (IVE) is one in which the user is perceptually surrounded by the VE. Ivan Sutherland (1965), one of the originators of 3-D computer graphics, was the first person to conceive and build an immersive VE system. For the history of IVEs, see Ellis (1995), Kalawsky (1993), and Rheingold (1991).

There are two usual implementations of an IVE. The first of these involves placing multiple projection screens and loudspeakers around the user. A popular design is the CAVE (Cruz-Neira, Sandin, & DeFanti, 1993), which involves back-projecting the computer-generated visual imagery onto the translucent walls, floor, and ceiling of a moderately sized cubical room, in which the user is free to move; shutter glasses provide stereoscopic stimulation, so that one sees the VE not as projections on the room surfaces, but as solid 3-D structures within and/or outside of the cube. The second and more common implementation of an IVE involves the use of a head-mounted display (HMD), used in conjunction with a computer and a head tracker (Barfield & Furness, 1995; Biocca & DeLaney, 1995; Burdea & Coffett, 1994; Durlach & Mavor, 1995; Kalawsky, 1993). The head tracker measures the changing position and orientation of the user's head within the physical environment, information that is communicated to the rendering computer, which has stored within it a 3-D representation of the simulated environment (Meyer, Applewhite, & Biocca, 1992). At any given moment, the computer generates and outputs the visual and auditory imagery to the user's HMD from a perspective that is based on the position and orientation of the user's head. The HMD consists of earphones and video displays attached to a support worn on the head; the video display component is based on cathode ray tube (CRT) displays, liquid crystal displays, or laser-based retinal scanners (Barfield, Hendrix, Bjornseth, Kaczmarek, &

With the support of Grant N00014-95-1-0573 from the Office of Naval Research (to J.M.L.) and National Science Foundation Grants SBR 9872084 (to J.J.B.) and SBR 9873432 (to J.M.L. and J.J.B.), and the authors have developed several immersive virtual displays, use of which has stimulated many of the ideas expressed here. The authors thank Florence Gaunet and Patrick Pélouch for comments on an earlier version of the article. Correspondence concerning this article should be addressed to J. M. Loomis, Department of Psychology, University of California, Santa Barbara, CA 93106 (e-mail: loomis@psych.ucsb.edu).

Related Work

Visualization of Geo-tagged Information

Related Work

Visualization of Geo-tagged Information

Jagan Sankaranarayanan[†]
jagan@cs.umd.edu

Benjamin E. Teitler[†]
bteitler@cs.umd.edu

Michael D. Lieberman[†]
codepoet@cs.umd.edu

Hanan Samet[†]
hjs@cs.umd.edu

Jon Spertling[‡]
Jon.Spertling@hud.gov

ABSTRACT

Twitter is an electronic medium that allows a large user population to communicate with each other simultaneously. Inherent to Twitter is an asymmetrical relationship between friends and followers that provides an interesting social network-like structure among the users of Twitter. Twitter messages, called tweets, are restricted to 140 characters and thus are usually very focused. We investigate the use of Twitter to build a news processing system, called *TwitterStand*, from Twitter tweets. The idea is to capture tweets that correspond to late breaking news. The result is that the identified distributed news wire service. The difference is that the identities of the contributors/reporters are not known in advance and there may be many of them. Furthermore, tweets are not sent according to a schedule: they occur as news is happening, and tend to be noisy while usually arriving at a high throughput rate. Some of the issues addressed include removing the noise, determining tweet clusters of interest bearing the relevant locations associated with the tweets.

Categories and Subject Descriptors

H.3 [Information Storage and Retrieval]: Information Storage and Retrieval

General Terms

Algorithms, Design, Performance

*This work was supported in part by the National Science Foundation under Grants EIA-08-12377, CCF-08-30618, and IIS-07-13501, as well as NVIDIA Corporation, Microsoft Research, Google, the E.T.S. Walton Visitor Award of the Science Foundation of Ireland, and the National Center for Geocomputation at the National University of Ireland at Maynooth.

[†]Department of Computer Science, Center for Automation Research, Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA.

[‡]HUD Office of Policy Development & Research (PD&R), 451 7th St. SW, Room 8146, Washington, DC 20410, USA.

Keywords

Twitter, News, Geotagging, Online clustering

1. INTRODUCTION

Twitter¹ is a social networking website that recently has been gaining much attention and following. Twitter is composed of users who send messages (termed *tweets*) to each other, where each tweet contains a maximum of 140 characters. At this time, it is estimated that there are 6 to 7 million users who use Twitter a total of 134 million times a month [4], and this number is increasing at a rapid rate. For example, for the year of 2008, Twitter grew in terms of the number of tweets sent at a rate of 1382% [12] which is a testament to the immense popularity and wide adoption of this service. The popularity of Twitter stems from its availability on a number of different electronic devices (e.g., web, cell phones, etc.), as well as the prevalence of a subculture of Twitter that encourages users to acquire a large friend pool, as well as send tweets on a wide variety of subjects, typically several times a day. The restriction on the lengths of Twitter messages invariably means that the tweets do not necessarily contain well formed ideas, being rather brief, yet complete enough so that users can make sense of the ideas that they convey. Note that tweets also have a mechanism by which the user can link to other objects on the web such as articles, images, videos, etc. (termed *artifacts*) which is typically used to link tweets to related material on the Internet.

The goal of this paper is to demonstrate how to use Twitter to automatically obtain breaking news from the tweets posted by Twitter users, and to provide a map interface for reading this news, since the geographic location of the user as well as the geographic terms comprising the tweets play an important role in *clustering* tweets and establishing clusters of geographic foci. In contrast to news aggregators such as Google News, Bing News, and Yahoo! News, we introduce a system called *TwitterStand* that works exclusively with only the tweets posted by the users of Twitter. The key novelty behind *TwitterStand* is one of mobilizing the millions of users in Twitter to be our eyes and ears in the world, bearing in mind that geographically proximate users often tweet about the same breaking news. In other words, we rely on Twitter users to be either providers of original news content (e.g., the 2008 Southern California earthquake [13] and the 2009 Iranian election [3]), or expressers of opinions on current news topics (i.e., mini blogs), both of which enable *TwitterStand* to automatically identify current news topics and cluster the corresponding tweets into appropriate news stories. We also associate an importance score with each news topic which can

Related Work

Visualization of Geo-tagged Information

Jagan Sankaranarayanan
jagan@cs.umd.edu

Benjamin E. Teitler
bteitler@cs.umd.edu

Michael D. Lieberman
codepoet@cs.umd.edu

Hanan Samet[†]
hjs@cs.umd.edu

Jon Sperling
Jon.Sperling@r

ABSTRACT

Twitter is an electronic medium that allows a large user population to communicate with each other simultaneously. Inherent to Twitter is an asymmetrical relationship between friends and followers that provides an interesting social network-like structure among the users of Twitter. Twitter messages, called tweets, are restricted to 140 characters and thus are usually very focused. We investigate the use of Twitter to build a news processing system, called *TwitterStand*, from usually very focused. The idea is to capture tweets that correspond to late breaking news. The result is that the distributed news wire service. The difference is that the contributors/reporters are not known in advance and there may be many of them. Furthermore, tweets are not sent according to a schedule: they occur as news is happening, and tend to be noisy while usually arriving at a high throughput rate. Some of the issues addressed include removing the noise, determining tweet clusters of interest bearing in mind that the methods must be online, and determining the relevant locations associated with the tweets.

Categories and Subject Descriptors

H.3 [Information Storage and Retrieval]: Information Storage and Retrieval

General Terms

Algorithms, Design, Performance

^{*}This work was supported in part by the National Science Foundation under Grants EIA-08-12377, CCF-08-30618, and IIS-07-13501, as well as NVIDIA Corporation, Microsoft Research, Google, the E.T.S. Walton Visitor Award of the Science Foundation of Ireland, and the National Center for Geocomputation at the National University of Ireland at Maynooth.

[†]Department of Computer Science, Center for Automation Research, Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA.

[‡]HUD Office of Policy Development & Research (PD&R), 451 7th St. SW, Room 8146, Washington, DC 20410, USA.

© 2009 Association for Computing Machinery. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or consultant of ACM. As such, the Government retains a certain right in this article, or to

Keywords

Twitter, News, Geotagging, Online

1. INTRODUCTION

Twitter¹ is a social networked been gaining much attention a posed of users who send mess other, where each tweet cont acters. At this time, it is es million users who use Twitt a month [4], and this numb For example, for the year o the number of tweets sent, a testament to the immense this service. The popular ability on a number of dif cell phones, etc.), as well Twitter that encourages as well as send tweets on several times a day. The messages invariably me contain well formed id enough so that users convey. Note that th the user can link to o images, videos, etc. f to link tweets to rel

The goal of this p ter to automaticall posted by Twitter reading this news as well as the ge an important rol ters' geographic Google News, B system called 7 the tweets pos behind Twitte in Twitter to mind that ge the same br users to be o 2008 South nian electic topics (i.e. to automa correspon associate

Placing Flickr Photos on a Map

Pavel Serdyukov^{*†}
Database Group
University of Twente
PO Box 217, 7500 AE
Enschede, The Netherlands
serdyukovpv@cs.utwente.nl

Vanessa Murdock
Yahoo! Research
Diagonal 177
08018 Barcelona, Spain
vmurdock@yahoo-inc.com

Roelof van Zwol
Yahoo! Research
Diagonal 177
08018 Barcelona, Spain
roelof@yahoo-inc.com

ABSTRACT

In this paper we investigate generic methods for placing photos uploaded to Flickr on the World map. As primary input for our methods we use the textual annotations provided by the users to predict the single most probable location where the image was taken. Central to our approach is a language model based entirely on the annotations provided by users. We define extensions to improve over the language model using tag-based smoothing and cell-based smoothing, and leveraging spatial ambiguity. Further we demonstrate how to incorporate GeoNames¹, a large external database of locations. For varying levels of granularity, we are able to place images on a map with at least twice the precision of the state-of-the-art reported in the literature.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]

General Terms

Algorithms, Measurement, Performance, Experimentation

Keywords

image localisation, language models, Flickr

1. INTRODUCTION

Due to the massive production of affordable GPS-enabled cameras and mobile phones [13, 16], location metadata such as *latitude* and *longitude* are automatically associated with the content generated by users. Users have the opportunity to spatially organise and browse their personal media, and photo sharing services are leading the growing enthusiasm for personal location-awareness [22]. Geo-referenced photos

^{*}Research performed while the author was an intern at Yahoo! Research.

[†]Also affiliated with TU Delft, ICT Group

¹<http://www.geonames.org> visited March 2009

can be organised in a browsable taxonomy of major locations or pin-pointed on a map to identify very small regions. Some of the most popular examples are Flickr Places² and Google Panoramio.³

While in theory every photo can be anchored to the location it was taken, in practice many photos are location agnostic. Furthermore, the majority of Flickr users do not own location-aware cameras. Thus a large proportion of photos uploaded to Flickr contain no location information even when the photo merits localizing. When uploading photos on Flickr users can still geo-tag their photos by dragging the photos to a particular point on the world map. This process is time-consuming and results in less accurate geo-tagging of photos compared to automatically geo-tagging photos from GPS-enabled cameras. When manually geo-tagging photos, Flickr initially suggests the location of the last uploaded photo or simply displays the world map.

The objective of this paper is to provide a more accurate starting point for geo-tagging photos, uploaded on Flickr, using the textual annotations provided by the user. According to recent literature [2, 21] users spend considerable effort to organise their “memory” geographically by describing photos with *tags* related to locations where they were taken. The location specific tags (such as *Torre Agbar* which is only located in Barcelona), and location related tags (such as *elephants* which are related to locations such as zoos, Africa and Asia) provide essential cues as to where a picture was taken. For photos that are location agnostic (such as *dog*), location information may or may not be provided, but it is normally not relevant to the context of the photo.

The literature related to geo-tagging of photos and its use is extensive. In particular the reverse problem of discovering important landmarks and events, given a geographic co-ordinate has been studied extensively [1, 17, 13]. However the problem of placing images on a map using the textual annotations provided by the user has received less attention. While we focus on Flickr as our primary data source, our approach

Related Work

Visualization of Geo-tagged Information

Jagan Sankaranarayanan
jagan@cs.umd.edu

Benjamin E. Teitler
bteitler@cs.umd.edu

Michael D. Lieberman
codepoet@cs.umd.edu

Hanan Samet
hjs@cs.umd.edu

Jon Sperling
Jon.Sperling@r

ABSTRACT

Twitter is an electronic medium that allows a large user population to communicate with each other simultaneously. Inherent to Twitter is an asymmetrical relationship between friends and followers that provides an interesting social network-like structure among the users of Twitter. Twitter messages, called tweets, are restricted to 140 characters and thus are usually very focused. We investigate the use of Twitter to build a news processing system, called *TwitterStand*, from Twitter tweets. The idea is to capture tweets that correspond to interesting news. The result is analogous to a

Keywords

Twitter, News, Geotagging, Online

1. INTRODUCTION

Twitter¹ is a social network that has been gaining much attention as a place where users who send messages to each other, where each tweet contains a maximum of 140 characters. At this time, it is estimated that over 10 million users who use Twitter

CHI 2008 Proceedings · Works In Progress

Category

H.3 [Information Systems] Storage

General Terms

Algorithms

*This work is supported by the National Science Foundation (NSF) under grant IIS-07-12345. Search, Information, and Visualization Research Group, University of Maryland, College Park, MD.

†Department of Computer Science, University of Maryland, College Park, MD.

‡HUD 07-12345, 7th St.

Content Visualization and Management of Geo-located Image Databases

Marco Cristani
Università degli Studi di Verona
Strada le Grazie 15
Verona, 37134 Italy
marco.cristani@univr.it

Alessandro Perina
Università degli Studi di Verona
Strada le Grazie 15
Verona, 37134 Italy
perina@sci.univr.it

Vittorio Murino
Università degli Studi di Verona
Strada le Grazie 15
Verona, 37134 Italy
Vittorio.murino@univr.it

Umberto Castellani
Università degli Studi di Verona

April 5-10, 2008 · Florence, Italy

Abstract

In the last years, several algorithms and platforms for photo sharing have been developed. Usually, in order to index huge quantities of images for a fast and intuitive retrieval, additional textual tags attached to the pictures are considered. In this paper, we present a set of solutions for an effective management of geo-located images, i.e. pictures equipped with tags indicating the geographical coordinates of acquisition. This brings towards an intuitive content visualization and management of large geo-located image databases.

Keywords

Image categorization, geo-located images, interfaces

Placing Flickr Photos on a Map

Pavel Serdyukov *
Database Group
University of Twente
PO Box 217, 7500 AE
Enschede, The Netherlands
serdyukovpv@cs.utwente.nl

Vanessa Murdock
Yahoo! Research
Diagonal 177
08018 Barcelona, Spain
vmurdock@yahoo-inc.com

Roelof van Zwol
Yahoo! Research
Diagonal 177
08018 Barcelona, Spain
roelof@yahoo-inc.com

ABSTRACT

In this paper we investigate generic methods for placing photos uploaded to Flickr on the World map. As primary input by the users to predict the single most probable location where the image was taken. Central to our approach is a language model based entirely on the text annotations provided by users. We define a

can be organised in a browsable taxonomy of major locations or pin-pointed on a map to identify very small regions. Some of the most popular examples are Flickr Places² and Google Panoramio.³

While in theory every photo can be anchored to the location it was taken, in practice many photos are location agnostic. Furthermore, the majority of Flickr users do not own geo-aware cameras. Thus a large proportion of photos uploaded to Flickr contain no location information even though the photo merits localizing. When uploading photos to a particular point on the world map, this process is often done by users can still geo-tag their photos by dragging the photo to a particular point on the world map. This process is often done manually and results in less accurate geo-tagging of photos compared to automatically geo-tagged photos from geo-aware cameras. When manually geo-tagging photos from geo-aware cameras, the user typically suggests the location of the last uploaded photo, which simply displays the world map.

The main objective of this paper is to provide a more accurate method for geo-tagging photos, uploaded on Flickr, using textual annotations provided by the user. According to literature [2, 21] users spend considerable effort to provide a “memory” geographically by describing photos related to locations where they were taken. The most common tags (such as *Torre Agbar* which is only in Barcelona), and location related tags (such as *Barcelona*) are related to locations such as zoos, Africa, and provide essential cues as to where a picture was taken. This information may or may not be provided, but it is often available related to geo-tagging of the photo.

In particular the reverse problem of discovering landmarks and events, given a geographic area studied extensively [1, 17, 13]. However, placing images on a map using the textual annotations by the user has received less attention. Flickr as our primary

Related Work

Visualization of Geo-tagged Information

Jagan Sankaranarayanan
jagan@cs.umd.edu

Benjamin E. Teitler[†]
bteitler@cs.umd.edu

Michael D. Lieberman[†]
codepoet@cs.umd.edu

Hanan Samet[†]
hjs@cs.umd.edu

Jon Sperling[†]
Jon.Sperling@h

ABSTRACT

Twitter is an electronic medium that allows a large user population to communicate with each other simultaneously. Inherent to Twitter is an asymmetrical relationship between friends and followers that provides an interesting social network-like structure among the users of Twitter. Twitter messages, called tweets, are restricted to 140 characters and thus are usually very focused. We investigate the use of Twitter to build a news processing system, called *TwitterStand*, from Twitter tweets. The idea is to capture tweets that correspond to interesting news. The result is analogous to a distributed database of tweets and their relationships, not sent in a single batch, opening, a throughpiping the in mind the relev

Category

H.3 [Inf Storage]

General

Algorithms

*This work is supported by the National Science Foundation (NSF) under grant IIS-07-11171. The authors would like to thank the anonymous reviewers for their helpful comments.

[†]Department of Computer Science, University of Maryland, College Park, MD 20742 USA
[‡]Department of Computer Science, University of Maryland, College Park, MD 20742 USA

CHI 2008 Proceedings - Works In Progress

TwitterStand: News in Tweets*

Keywords

Twitter, News, Geotagging, Online

1. INTRODUCTION

Twitter¹ is a social networked service that has been gaining much attention among users who send messages to each other, where each tweet contains at most 140 characters. At this time, it is estimated that there are over 10 million users who use Twitter.

Content Visualization of Geo-located Images

Marco Cristani
Università degli Studi di Verona
Strada le Grazie 15
Verona, 37134 Italy
marco.cristani@univr.it

Alessandro Perina
Università degli Studi di Verona
Strada le Grazie 15
Verona, 37134 Italy
perina@sci.univr.it

Umberto Castellani
Università degli Studi di Verona

Vittorio Murino
Università degli Studi di Verona
Strada le Grazie 15
Verona, 37134 Italy
Vittorio.murino@univr.it

Abstract

In this paper we investigate the problem of visualizing geo-located images. We propose a system that indexes and visualizes the content of geo-located images. This is done by indexing the images and their metadata.

Key

Images

ABSTRACT

In this paper we investigate the problem of visualizing geo-located images. We propose a system that indexes and visualizes the content of geo-located images. This is done by indexing the images and their metadata.

Placing Flickr Photos on a Map

Pavel Serdyukov^{*†}
Database Group
University of Twente
PO Box 217, 7500 AE
Enschede, The Netherlands
serdyukovpv@cs.utwente.nl

Vanessa Murdock
Yahoo! Research
Diagonal 177
08018 Barcelona, Spain
vmurdock@yahoo-inc.com

Roelof van Zwol
Yahoo! Research
Diagonal 177
08018 Barcelona, Spain
roelof@yahoo-inc.com

PhotoStand: A Map Query Interface for a Database of News Photos

Hanan Samet Marco D. Adelfio Brendan C. Fruin
Michael D. Lieberman Jagan Sankaranarayanan
Center for Automation Research, Institute for Advanced Studies,
Department of Computer Science, University of Maryland
College Park, MD 20742 USA
{hjs, marco, brendan, codepoet, jagan}@cs.umd.edu

ABSTRACT

PhotoStand enables the use of a map query interface to retrieve news photos associated with news articles that are in turn associated with the principal locations that they mention collected as a result of monitoring the output of over 10,000 RSS news feeds, made available within minutes of publication, and stored in a PostgreSQL database. The news photos are ranked according to their relevance to the clusters of news articles associated with locations at which they are displayed. This work differs from traditional work in this field as the associated locations and topics (by virtue of the cluster with which the articles containing the news photos are associated) are generated automatically without any human intervention such as tagging, and that photos are retrieved by location instead of just by keyword as is the case for many existing systems. In addition, the clusters provide a filtering step for detecting near-duplicate news photos.

1. INTRODUCTION

A demo is presented of PhotoStand (see also the related NewsStand [9, 17, 21, 29], TwitterStand [6, 24], and STEWARD [12] systems) which is an example application of a general framework we are developing for retrieving multimedia data (e.g., text, images, videos) using a map query interface from a database of news articles, photos, and videos (i.e., by location in real-time which differentiates it from other systems).

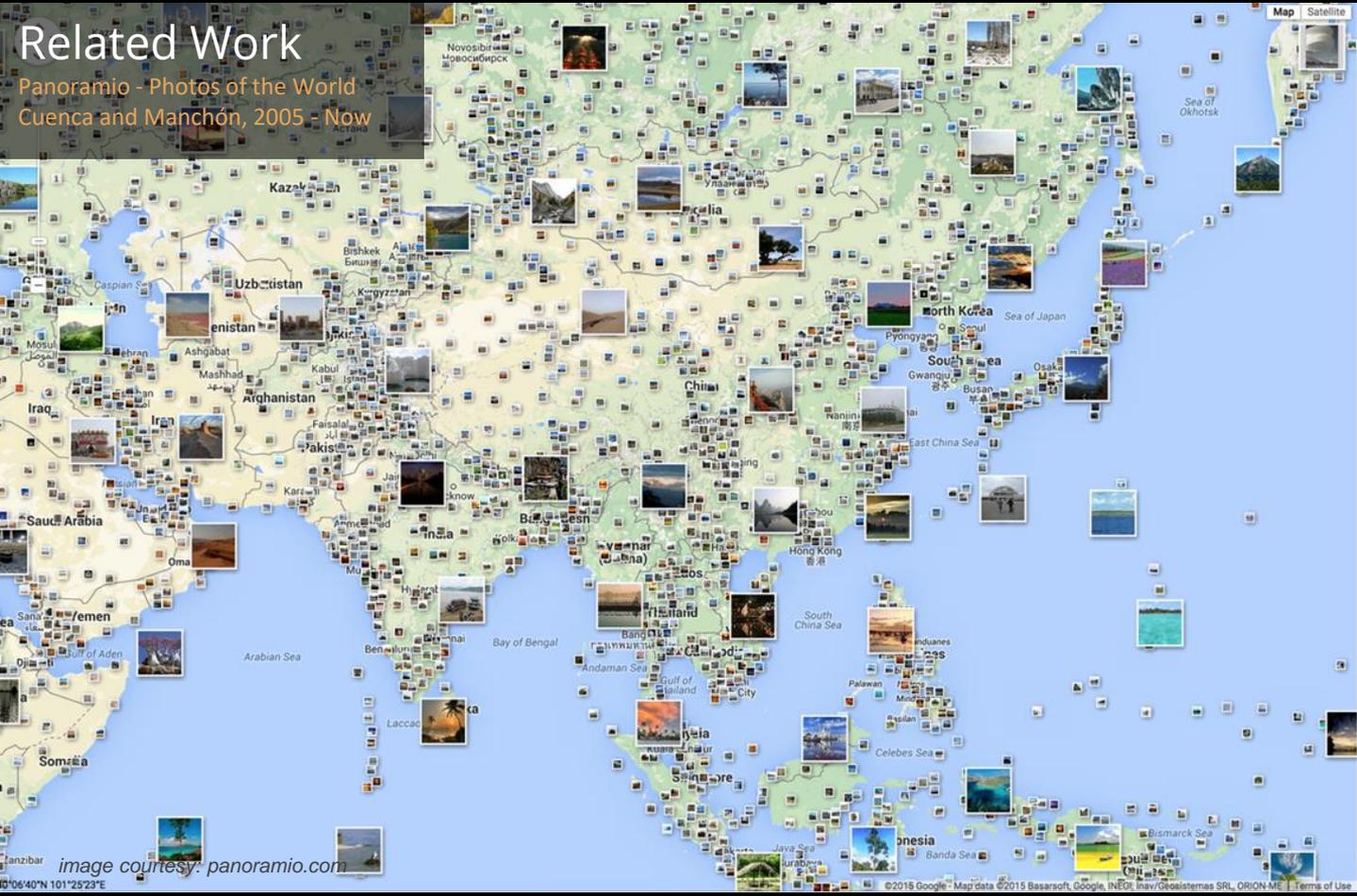
articles, enabling them to be accessed by spatial queries such as windowing or simple point location; and its *clusterer* [30], which groups articles about the same topic. A key to the NewsStand database system is its pipe server which coordinates its processing modules by assigning batches of articles to them. NewsStand's user interface enables the retrieval of clusters of news articles for display using its map user interface by executing what we term *top-k window queries*. At present, NewsStand database of articles currently containing about 300GB of data.

The PhotoStand and TweetPhoto [3] demos are related in the sense that PhotoStand uses photos from news articles in NewsStand, while TweetPhoto uses photos from news tweets in TwitterStand [24]. In addition, the PhotoStand demo demonstrates the database querying capability of NewsStand as well as its capability to do similarity searching for news photos where the first step in the similarity detection process is based on the text associated with the photos, while the second step involves use of the actual image features (e.g., texture, color) to enable detecting near duplicates, thereby avoiding the combinatorial complexity of comparing every photo with every other photo.

The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 indicates how news articles (and consequently news photos) are processed.

Related Work

Panoramio - Photos of the World
Cuenca and Manchón, 2005 - Now
Acuña



Map Satellite

Popular Recent Places Indoor

Also show photos not selected for Google Earth

©2015 Google - Map data ©2015 Basarsoft, Google, INEGI, Inav/Geosistemas SRL, ORION-ME | Terms of Use

Popular photos in Google Earth

« Previous Next »

image courtesy: panoramio.com

Related Work

TwitterStand: News in Tweets
Sankaranarayanan et al. SIGGIS 2009

Top Stories Map Mode Time Mode

Enter a location...

Keyword Search...

Display: 50

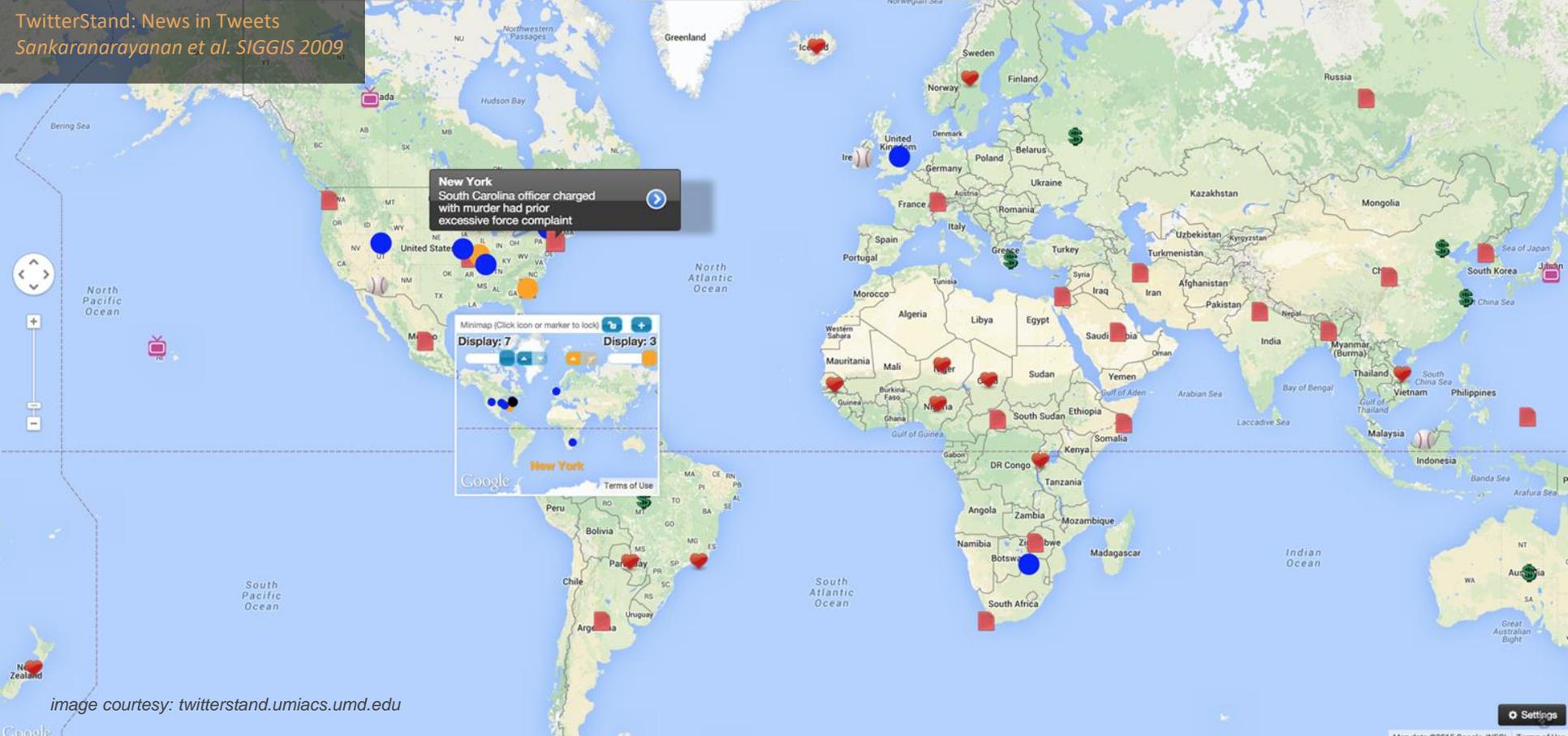
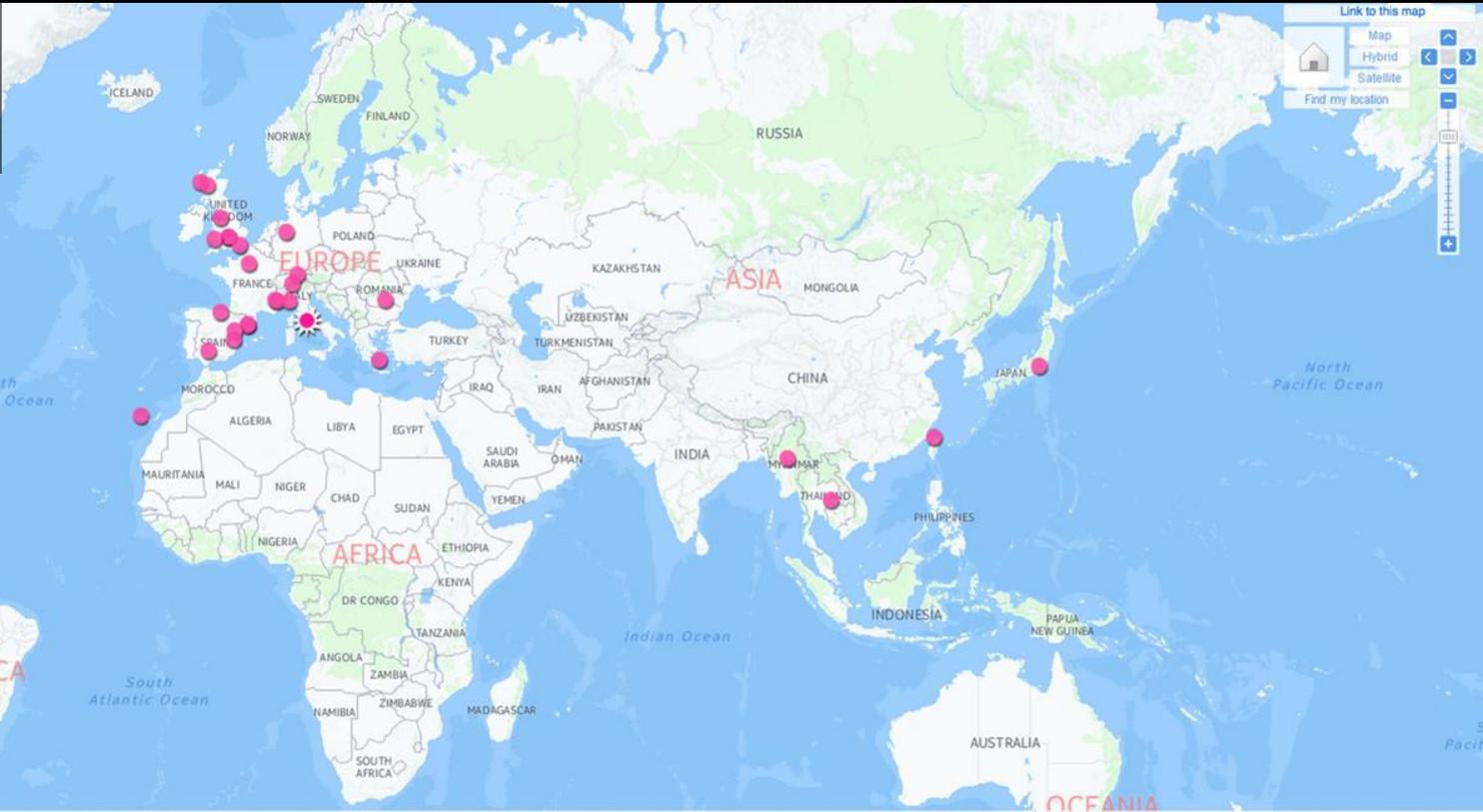


image courtesy: twitterstand.umiacs.umd.edu

Related Work

Flickr - World Map
Serdyukov et al. SIGIR 2009



2,175,866 geotagged items
Sort by: Interesting • Recent

Search the map

Search Go

Locate Go

92

Bering Sea

Sea of Okhotsk

North Pacific Ocean

Bismarck Sea

- All
- General
- Business
- SciTech
- Entertainment
- Health
- Sports

Related Work

PhotoStand for News Photos
Samet et al. VLDB 2013

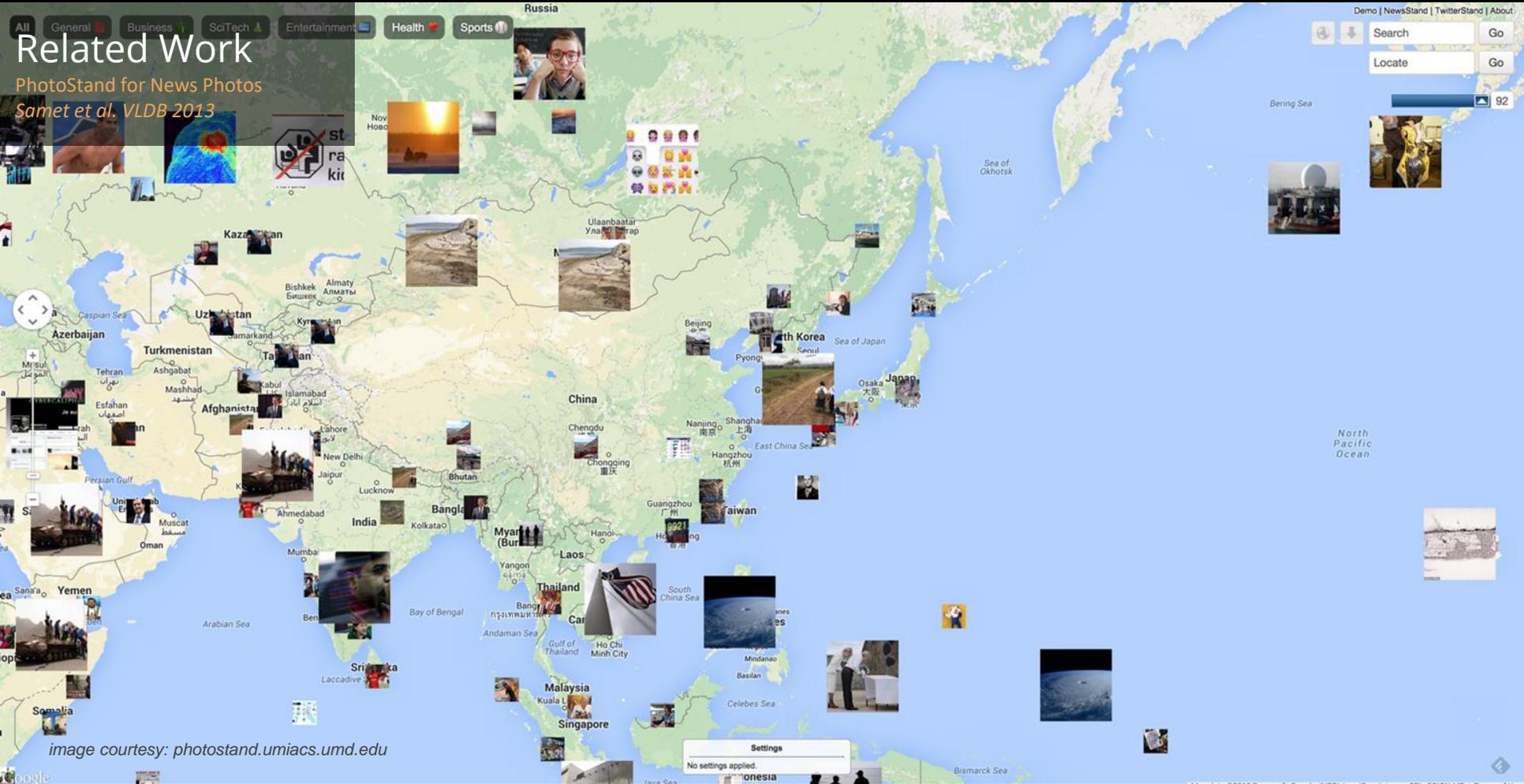


image courtesy: photostand.umiacs.umd.edu

Related Work

*Visualization of Geo-tagged Information
(cont.)*

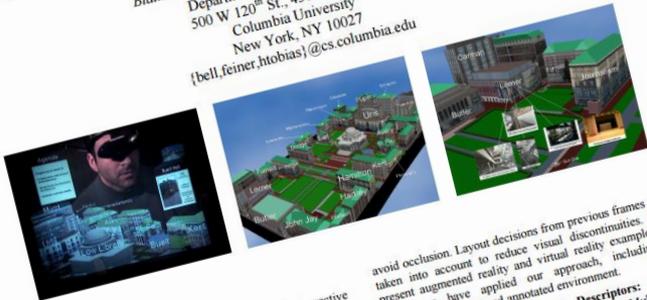
Related Work

Visualization of Geo-tagged Information (cont.)

UIST 2001 (ACM Symp. on User Interface Software and Technology), Orlando, FL, November 11-14, 2001, pp. 101-110

View Management for Virtual and Augmented Reality

Blaine Bell Steven Feiner Tobias Hollerer
Department of Computer Science
500 W 120th St., 450 CS Building
Columbia University
New York, NY 10027
{bell,feiner,tobias}@cs.columbia.edu



ABSTRACT

We describe a view-management component for interactive 3D user interfaces. By *view management*, we mean maintaining visual constraints on the projections of objects on the view plane, such as locating related objects by other, or preventing objects from occluding each other. Our view-management component accomplishes this by modifying selected object properties, including position, size, and transparency, which are tagged to indicate their constraints. For example, some objects may have geometric properties that are determined entirely by a physical simulation and which cannot be modified, while other objects may be annotations whose position and size are flexible.

We introduce algorithms that use upright rectangular extents to represent on the view plane a dynamic and efficient approximation of the occupied space containing the projections of visible portions of 3D objects, as well as the unoccupied space in which objects can be placed to

avoid occlusion. Layout decisions from previous frames are taken into account to reduce visual discontinuities. We present augmented reality and virtual reality examples to which we have applied our approach, including a dynamically labeled and annotated environment.

CR Categories and Subject Descriptors: H.5.1 [Information Interfaces and Presentation] Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation] User Interfaces—Graphical User Interfaces, Screen design; I.3.6 [Computer Graphics] Methodology and Techniques—Interaction Techniques; I.3.7 [Computer Graphics] Three-Dimensional Graphics and Realism—Virtual Reality.

Additional Keywords and Phrases: view management, environment management, annotation, labeling, wearable computing, augmented reality, virtual environments

1. INTRODUCTION

Designing a 3D graphical user interface (UI) requires creating a set of objects and their properties, arranging them in a scene, setting a viewing specification, and deciding how to light and rendering parameters, and deciding how to update these decisions for each frame. Some of these decisions may be fully constrained, for example, a scene's viewing specification may be explicitly constrained by a viewing specification that is determined by a viewing specification, other decisions must be made, especially interested in

Related Work

Visualization of Geo-tagged Information
(cont.)

UIST 2001 (ACM Symp. on User Interface Technology)

View Management



ABSTRACT
We describe a view-management system for 3D user interfaces. By maintaining visual constraints on the view plane, such as other, or preventing object view-management constraints, size, and transparency. For example, properties that are a simulation and which objects may be annotated flexibly.

We introduce algorithms that extend to represent efficient approximations of the projections of the unoccupied space.

Photo Tourism: Exploring Photo Collections in 3D

Noah Snavely
University of Washington

Steven M. Seitz
University of Washington

Richard Szeliski
Microsoft Research

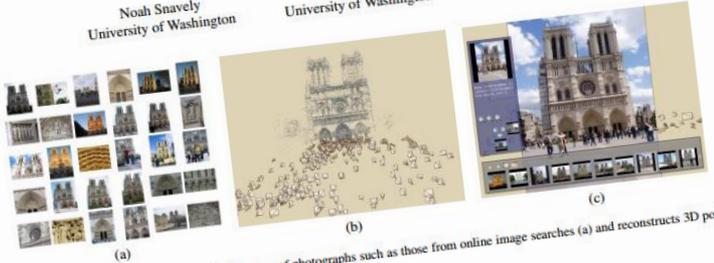


Figure 1: Our system takes unstructured collections of photographs such as those from online image searches (a) and reconstructs 3D points and viewpoints (b) to enable novel ways of browsing the photos (c).

Abstract

We present a system for interactively browsing and exploring large unstructured collections of photographs of a scene using a novel 3D interface. Our system consists of an image-based modeling front end that automatically computes the viewpoint of each photograph as well as a sparse 3D model of the scene and image to model correspondences. Our *photo explorer* uses image-based rendering techniques to smoothly transition between the set of images and enabling full 3D navigation and exploration of the set of images and world geometry, along with auxiliary information such as overhead maps. Our system also makes it easy to construct photo tours of scenic or historic locations, and to annotate image details, which are automatically transferred to other relevant images. We demonstrate our system on several large personal photo collections as well as images gathered from Internet photo sharing sites.

CR Categories: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities 1.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Modeling and recovery of physical attributes

Keywords: image-based rendering, image-based modeling, photo browsing, structure from motion

1 Introduction

One goal of image-based rendering is to evoke a visceral sense of a scene from a collection of photographs of a scene. The tool

is that these approaches will one day allow virtual tourism of the world's most interesting and important sites.

During this same time, digital photography, together with the Internet, have combined to enable sharing of photographs on a truly massive scale. For example, a Google image search on “Notre Dame Cathedral” returns over 15,000 photos, capturing the scene from myriad viewpoints, levels of detail, lighting conditions, seasons, decades, and so forth. Unfortunately, the proliferation of shared photographs has outpaced the technology for browsing such collections, as tools like Google (www.google.com) and Flickr (www.flickr.com) return pages and pages of thumbnails that the user must comb through.

In this paper, we present a system for browsing and organizing large photo collections of popular sites which exploits the common 3D geometry of the underlying scene. Our approach is based on computer graphics themselves, the photographers’ localizations and orientations, along with a sparse 3D geometric representation of the scene, using a state-of-the-art image-based modeling system. Our system handles large collections of unorganized photographs taken by different cameras in widely different conditions. We show how the inferred camera and scene information enables the following capabilities:

- **Scene visualization.** Fly around popular world sites in 3D by morphing between photos.
- **Object-based photo browsing.** Show me more images that contain this object or part of the scene.
- **Where was I?** Tell me where I was when I took this picture.
- **What am I looking at?** Tell me about objects visible in this image by transferring annotations from similar images.

Related Work

Visualization of Geo-tagged Information
(cont.)

UIST 2001 (ACM Symp. on User Interface Support)

View Management



ABSTRACT

We describe a view-management system for 3D user interfaces. By maintaining visual constraints on the view plane, such as other, or preventing object view-management constraints, modifying selected object size, and transparency. For example, constraints that are a properties that are a simulation and which objects may be annotated.

We introduce algorithms to represent efficient approximations of the projections of the unoccupied space.

Photo Tourism: Enabling

Noah Snavely
University of Washington



(a)

Figure 1: Our system takes unstructured collections of photographs and viewpoints (b) to enable novel ways of

Abstract

We present a system for interactively browsing unstructured collections of photographs in a 3D interface. Our system consists of a front end that automatically computes a 3D model of the scene as well as a sparse 3D model of the correspondences. Our *photo explorer* techniques to smoothly transition between enabling full 3D navigation and exploring world geometry, along with auxiliary maps. Our system also makes it easy to scenic or historic locations, and to are automatically transferred to other. We illustrate our system on several large photo collections gathered from Internet photo

CR Categories: H.5.1 [Informal Multimedia Information Systems] Multimedia realities I.2.10 [Artificial Intelligence]—Modeling and Understanding

Keywords: image-based rendering, browsing, structure from motion

1 Introduction

Social Snapshot: A System for Temporally Coupled Social Photography

Robert Patro, Cheuk Yiu Ip, Sujal Bista, and Amitabh Varshney • University of Maryland, College Park

Since the invention of photography, taking pictures of people, places, and activities has become integral to our lives. In the past, only purposeful, precious moments were the primary subjects of photography. But technological advances have brought photography to our everyday lives in the form of compact cameras and even cell phone cameras.

Social Snapshot actively acquires and reconstructs temporally dynamic data. The system enables spatiotemporal 3D photography using commodity devices, assisted by their auxiliary sensors and network functionality. It engages users, making them active rather than passive participants in data acquisition.

struments used to acquire photographs are tediously calibrated to produce precise measurements.

To simplify 3D photography, our Social Snapshot system performs active acquisition and reconstruction of temporally dynamic data. Using multiple users' cell phone cameras and no preliminary calibration, it achieves approximate but visually convincing renderings of 3D scenes, even though

Social Snapshot's Contributions

Social Snapshot's contributions fit naturally into two categories: technical and social.

The technical contributions are improved algorithms and techniques that enhance our system's novelty and scalability. For example, Social Snapshot produces a textured and colored-mesh reconstruction from a loosely ordered photo collection, rather than the sparse or dense point reconstructions produced by related approaches. In addition, it features locally optimized mesh generation and viewing. Finally, it provides camera network capabilities to support synchronized capture of temporally dynamic data.

The social contributions lead to a new way of thinking about the interplay between data acquisition and social interactions. They also let us define social photography as an active, rather than a passive, endeavor. For example, Social Snapshot encourages collaborative photography as a social endeavor, letting users capture dynamic action by synchronizing their photographs. It leverages social trends such as online media sharing and event organization to spur a novel data acquisition mode.

For a look at some of the previous research on which Social Snapshot is based, see the "Related Work in Scene Visualization and Computer Vision" sidebar on pages 78–79.

Related Work

Visualization of Geo-tagged Information
(cont.)

UIST 2001 (ACM Symp. on User Interface Research)

Photo Tourism: Enabling Novel Views of User-generated Photographs

Noah Snavely
University of Washington



(a)



(b)

Figure 1: Our system takes unstructured collections of photographs and viewpoints (b) to enable novel ways of viewing the scene.

Abstract

We present a system for interactively browsing unstructured collections of photographs in a 3D interface. Our system consists of a front end that automatically computes a scene as well as a sparse 3D model of the scene. Our *photo explorer* techniques to smoothly transition between world geometry, along with auxiliary maps. Our system also makes it easy to explore or historic locations, and to transfer our system on several large public images gathered from Internet photo sharing sites.

CR Categories: H.5.1 [Informal Multimedia Information Systems]—Virtual realities I.2.10 [Artificial Intelligence]—Modeling and Simulation

Keywords: image-based rendering, structure from motion

1 Introduction

The explosion of image-based content on the Web has led to a proliferation of image-based user interfaces. By maintaining visual context in the view plane, such as other, or preventing object view-management constraints. For example, size, and transparency, which are properties that are difficult to simulate and which objects may be annotated flexibly.

We introduce algorithms to represent efficient approximations of the projections of the unoccupied space.

Social A Syst Social

Robert Patro, Cl

3D Wikipedia: Using online text to automatically label and navigate reconstructed geometry

Bryan C. Russell¹ Ricardo Martin-Brualla² Daniel J. Butler² Steven M. Seitz² Luke Zettlemoyer²
¹Intel Labs ²University of Washington



Figure 1: Given a reference text describing a specific site, for example the Wikipedia article above for the Pantheon, we automatically create a labeled 3D reconstruction, with objects in the model linked to where they are mentioned in the text. The user interface enables coordinated browsing of the text with the visualization (see video).

Abstract

We introduce an approach for analyzing Wikipedia and other text together with online photos, to produce *annotated* 3D models of famous tourist sites. The approach is completely automated, and leverages online text and photo co-occurrences via Google Image Search. It enables a number of new interactions, which we demonstrate in a new 3D visualization tool. Text can be selected to move the camera to the corresponding objects, 3D bounding boxes provide anchors back to the text describing them, and the overall narrative of the text provides a temporal guide for automatically flying through the scene to visualize the world as you read about it. We show compelling results on several major tourist sites.

CR Categories: H.5.1 [Information Interfaces and Presentation]—Multimedia Information Systems—Artificial Intelligence; and Processing—Text analysis I.2.7 [Artificial Intelligence]; Natural Language and Scene Understanding—Modeling and recovery of physical attributes

Keywords: image-based modeling and rendering, Wikipedia, natural language processing, 3D visualization

Links: [DL](#) [PDF](#)

1 Introduction

Tourists have long relied on guidebooks and other reference texts to learn about and navigate sites of interest. While guidebooks are packed with interesting historical facts and descriptions of scenes they present, it can be difficult to fully visualize the with the text, but coverage is sparse and it can be difficult to understand the spatial relationships between each image viewpoint. For example, the Berlitz and Lonely Planet guides [Berlitz International 2003; Garwood and Hole 2012] for Rome each contain just a single photo of the Pantheon, and have a similar lack of photographic coverage of other sites. Even online sites such as Wikipedia, which do not have space restrictions, have similarly sparse and disconnected visual coverage.

Instead of relying exclusively on static images embedded in text, suppose you could create an interactive, photorealistic visualization, where, for example, a Wikipedia page is shown next to a detailed 3D model of the described site. When you select an object in the scene via a smooth, photorealistic transition. Similarly, when you click on an object in the visualization, it highlights the corresponding descriptive text on the Wikipedia page. Our goal is to create such a visualization **completely automatically** by analyzing the Wikipedia page itself, together with many photos available online (Figure 1).

Our Approach?

Social Street View

Social Street View



*An immersive social media
navigation system in mixed-reality!*

Demonstration

The Augmentarium, UMIACS
6000 x 3000 pixels



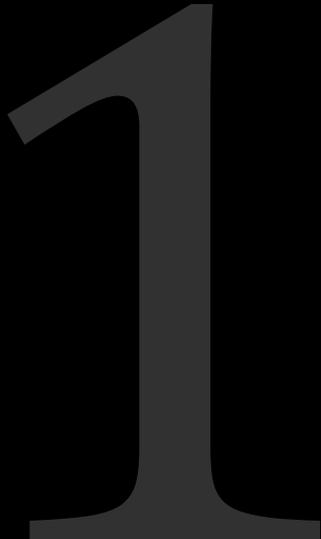
Natural Immersive Virtual Reality?

almost never in a natural immersive virtual reality settings.

Conception, architecting & implementation

Social Street View

A mixed reality system that can depict geo-tagged social media in immersive 3D web environments



Blending multiple modalities of

2

Street View + Social Media

Depth maps, normal maps, and road orientation
GPS coordinates and time creation

Enhancing visual augmentation

3

Maximal Poisson-disk sampling

Evaluated by image saliency metrics

Achieving cross-platform compatibility by

4

WebGL + Three.js

smartphones, tablets, desktop, high-resolution
large-area wide field of view tiled display walls, as
well as head-mounted displays.

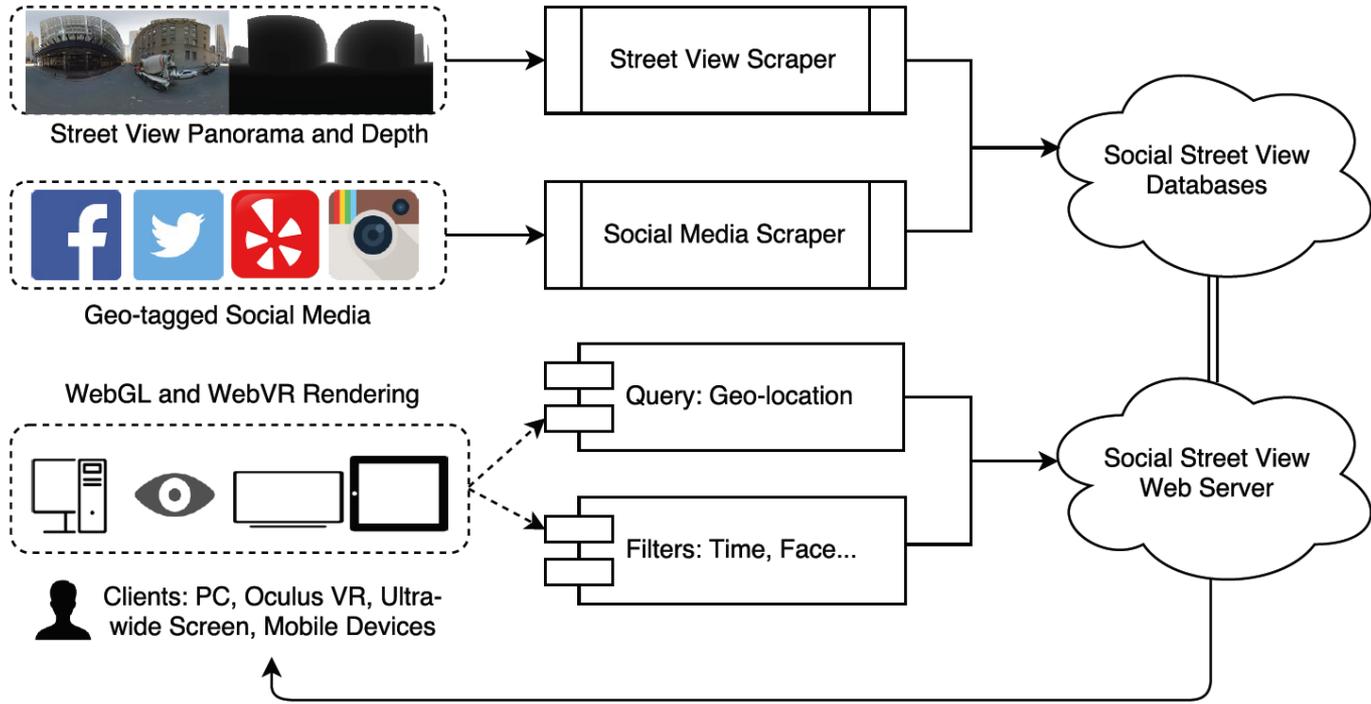
Technical Challenges?



System Overview

Architecture

Social Street View System Flowchart





Street View Cars - Cameras, Lasers and GPS

Image courtesy from Google Street View







Tiles of Panoramic Images

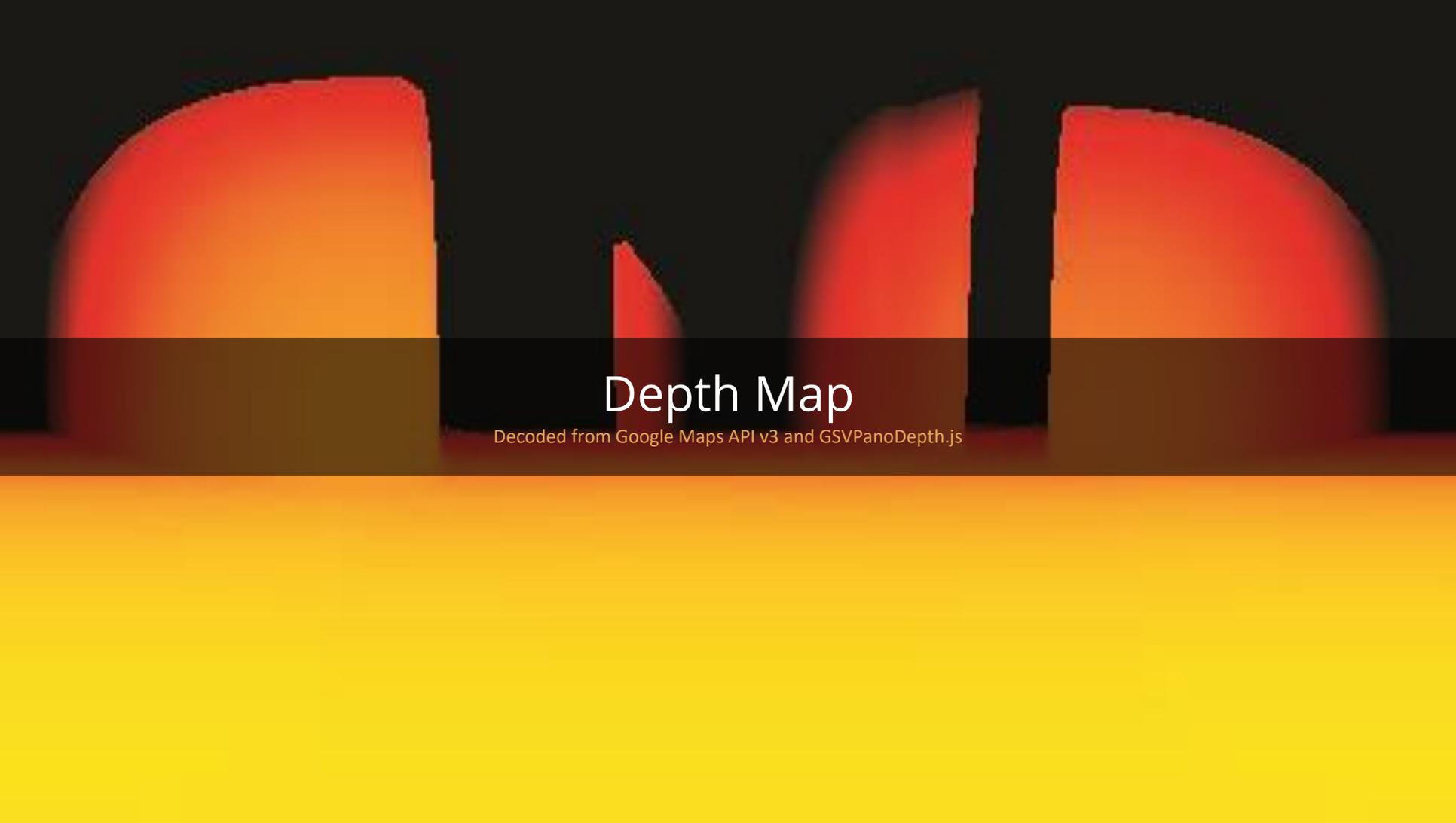
Image courtesy from Google Street View





Panorama

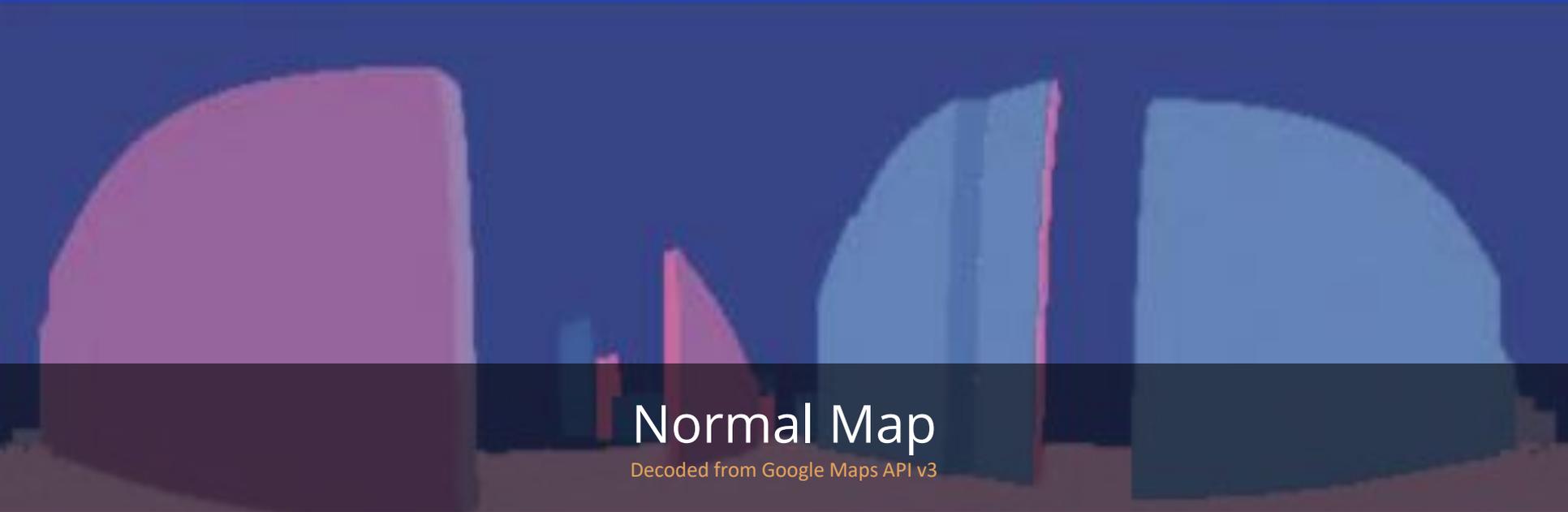




Depth Map

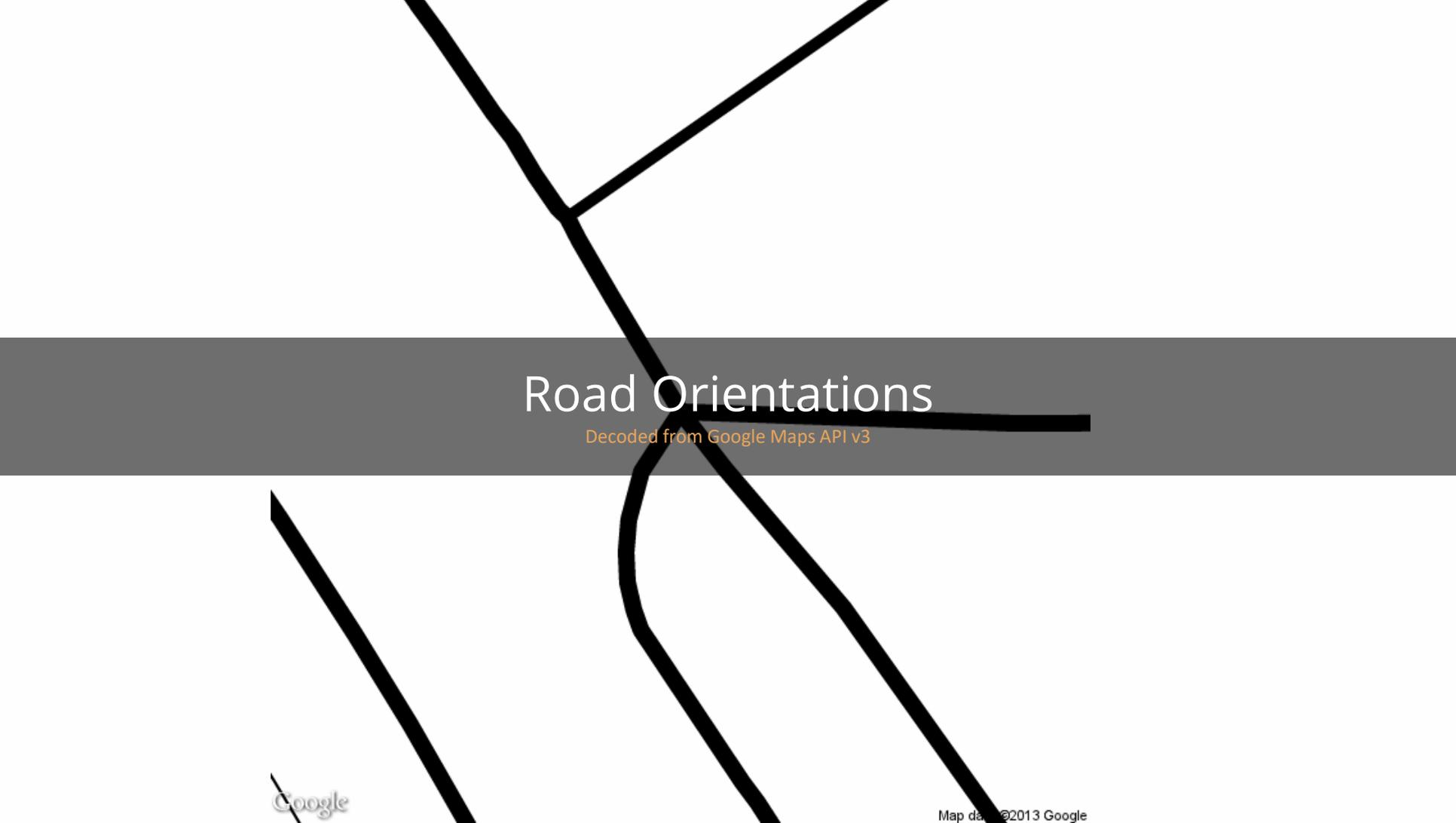
Decoded from Google Maps API v3 and GSVPanoDepth.js





Normal Map

Decoded from Google Maps API v3



Road Orientations

Decoded from Google Maps API v3

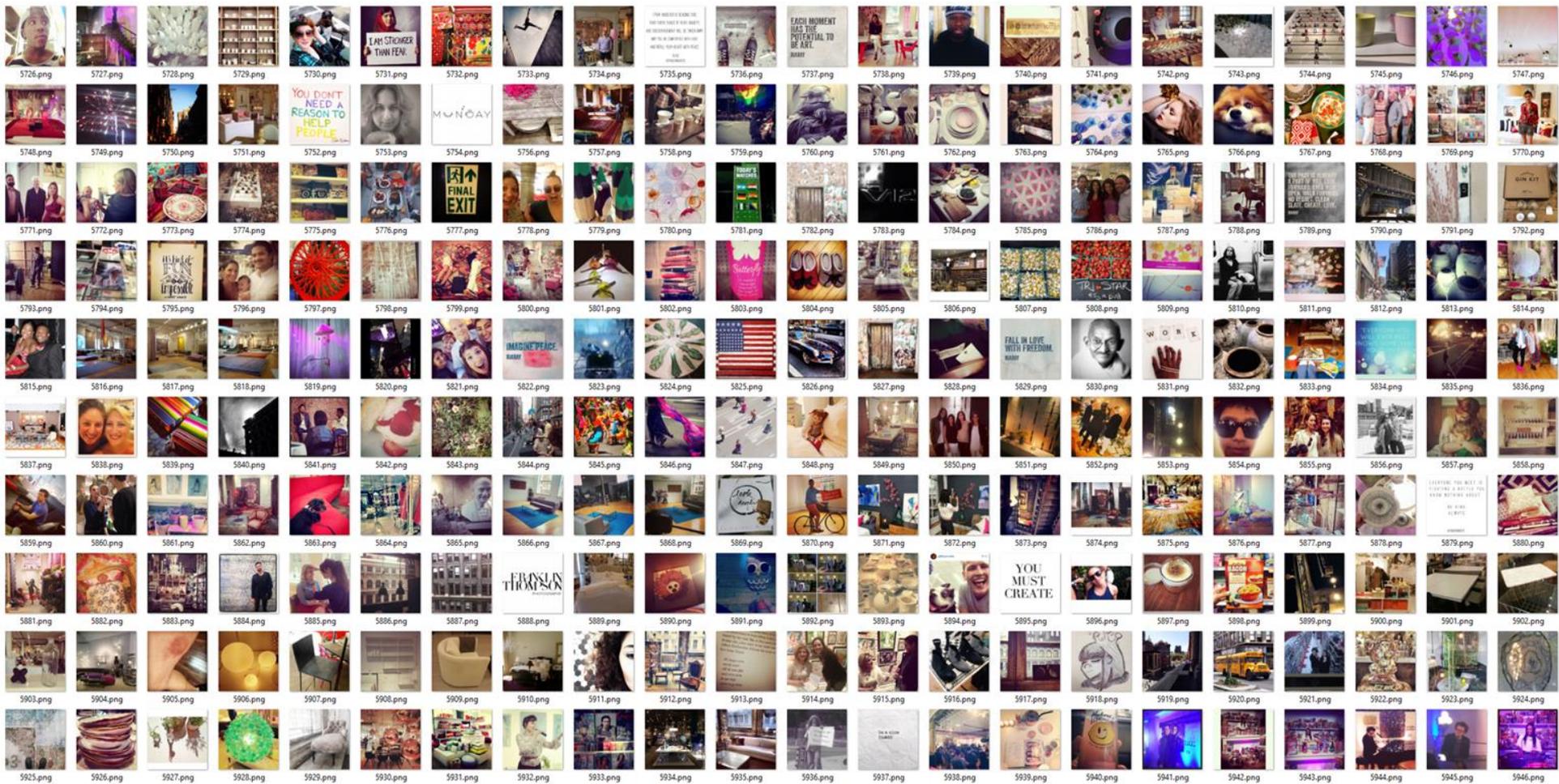
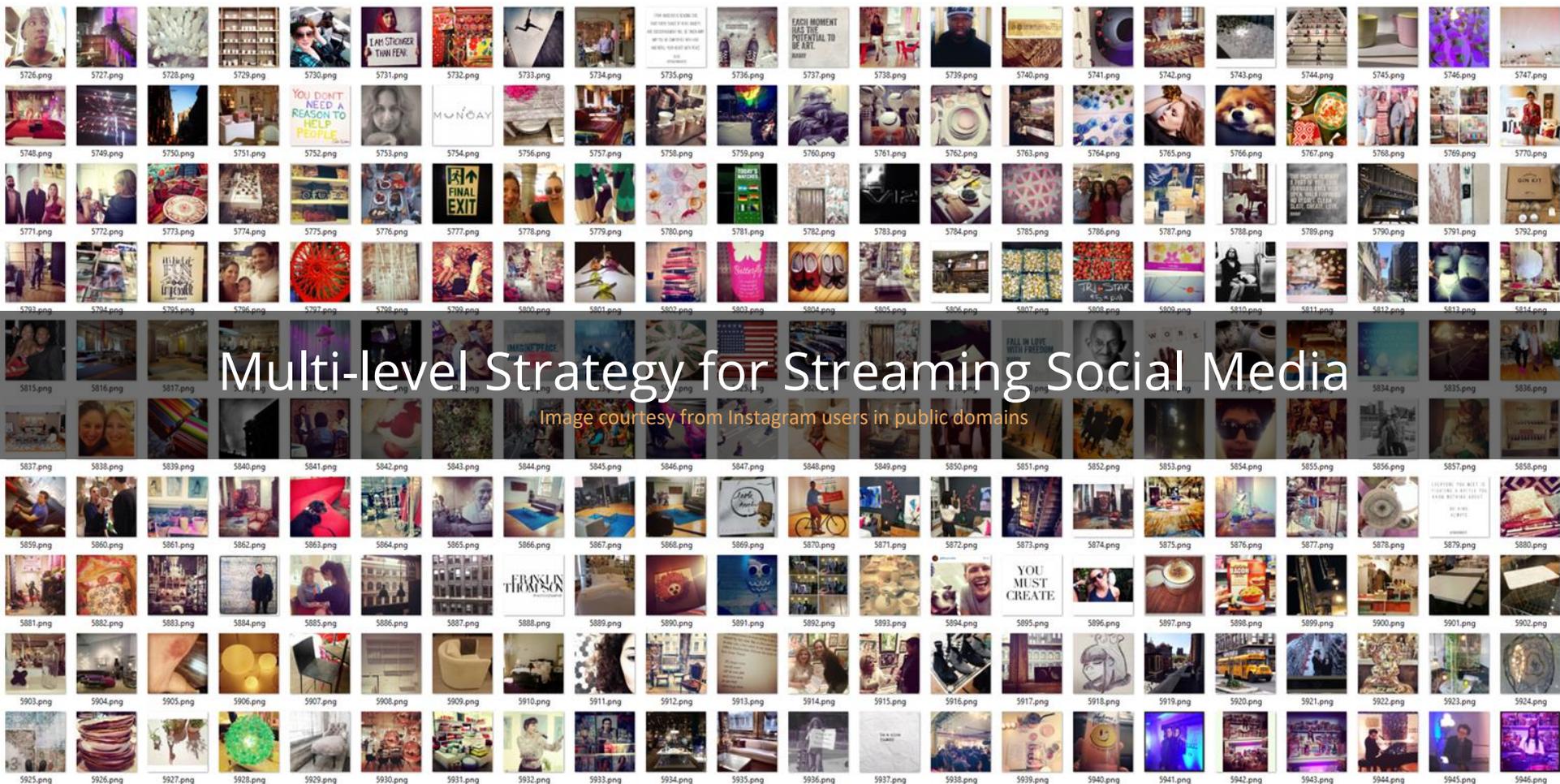


image courtesy: Instagram



Multi-level Strategy for Streaming Social Media

Image courtesy of Instagram users in public domains

Bipartite Graph

$$G = V(V, S, E)$$



Haversine Formula

Andrew, 1805

$$\alpha_{ij} = \sin^2\left(\frac{\varphi_i - \varphi_j}{2}\right) + \cos \varphi_i \cdot \cos \varphi_j \cdot \sin^2\left(\frac{\lambda_i - \lambda_j}{2}\right)$$

$$\beta_{ij} = 2 \cdot \operatorname{atan2}\left(\sqrt{\alpha_{ij}}, \sqrt{(1 - \alpha_{ij})}\right)$$

$$d_{ij} = R \cdot \beta_{ij}$$

Interface

Keywords Search

Urban Rural Indoor Terrain Misc

NYC London Paris Rome Tokyo

Month 1 - 9

Hour 3 - 20

Distance 0 - 64

2D Depth Side Front Model

Enhance Dark None Enhance Bright

Spring Summer Autumn Winter

All Face-only None-face-only

All No-text-only With-text-only Include-video

DirLight Unlimited

AmbLight Unlimited

BloomEffect Unlimited

DarkEffect Unlimited

LightColor Unlimited

Number Unlimited

Radius Unlimited

Help

Credits



Map | Satellite

W 50th St

W 49th St

W 48th St

W 47th St

W 46th St

W 45th St

W 44th St

W 43rd St

W 42nd St

W 41st St

W 40th St

W 39th St

W 38th St

W 37th St

W 36th St

W 35th St

W 34th St

W 33rd St

W 32nd St

W 31st St

W 30th St

W 29th St

W 28th St

W 27th St

W 26th St

W 25th St

W 24th St

W 23rd St

W 22nd St

W 21st St

W 20th St

W 19th St

W 18th St

W 17th St

W 16th St

W 15th St

W 14th St

W 13th St

W 12th St

W 11th St

W 10th St

W 9th St

W 8th St

W 7th St

W 6th St

W 5th St

W 4th St

W 3rd St

W 2nd St

W 1st St

W 0th St

W -1st St

W -2nd St

W -3rd St

W -4th St

W -5th St

W -6th St

W -7th St

W -8th St

W -9th St

W -10th St

W -11th St

W -12th St

W -13th St

W -14th St

W -15th St

W -16th St

W -17th St

W -18th St

W -19th St

W -20th St

W -21st St

W -22nd St

W -23rd St

W -24th St

W -25th St

W -26th St

W -27th St

W -28th St

W -29th St

W -30th St

W -31st St

W -32nd St

W -33rd St

W -34th St

W -35th St

W -36th St

W -37th St

W -38th St

W -39th St

W -40th St

W -41st St

W -42nd St

W -43rd St

W -44th St

W -45th St

W -46th St

W -47th St

W -48th St

W -49th St

W -50th St

W -51st St

W -52nd St

W -53rd St

W -54th St

W -55th St

W -56th St

W -57th St

W -58th St

W -59th St

W -60th St

W -61st St

W -62nd St

W -63rd St

W -64th St

W -65th St

W -66th St

W -67th St

W -68th St

W -69th St

W -70th St

W -71st St

W -72nd St

W -73rd St

W -74th St

W -75th St

W -76th St

W -77th St

W -78th St

W -79th St

W -80th St

W -81st St

W -82nd St

W -83rd St

W -84th St

W -85th St

W -86th St

W -87th St

W -88th St

W -89th St

W -90th St

W -91st St

W -92nd St

W -93rd St

W -94th St

W -95th St

W -96th St

W -97th St

W -98th St

W -99th St

W -100th St

W -101st St

W -102nd St

W -103rd St

W -104th St

W -105th St

W -106th St

W -107th St

W -108th St

W -109th St

W -110th St

W -111st St

W -112nd St

W -113rd St

W -114th St

W -115th St

W -116th St

W -117th St

W -118th St

W -119th St

W -120th St

W -121st St

W -122nd St

W -123rd St

W -124th St

W -125th St

W -126th St

W -127th St

W -128th St

W -129th St

W -130th St

W -131st St

W -132nd St

W -133rd St

W -134th St

W -135th St

W -136th St

W -137th St

W -138th St

W -139th St

W -140th St

W -141st St

W -142nd St

W -143rd St

W -144th St

W -145th St

W -146th St

W -147th St

W -148th St

W -149th St

W -150th St

W -151st St

W -152nd St

W -153rd St

W -154th St

W -155th St

W -156th St

W -157th St

W -158th St

W -159th St

W -160th St

W -161st St

W -162nd St

W -163rd St

W -164th St

W -165th St

W -166th St

W -167th St

W -168th St

W -169th St

W -170th St

W -171st St

W -172nd St

W -173rd St

W -174th St

W -175th St

W -176th St

W -177th St

W -178th St

W -179th St

W -180th St

W -181st St

W -182nd St

W -183rd St

W -184th St

W -185th St

W -186th St

W -187th St

W -188th St

W -189th St

W -190th St

W -191st St

W -192nd St

W -193rd St

W -194th St

W -195th St

W -196th St

W -197th St

W -198th St

W -199th St

W -200th St

W -201st St

W -202nd St

W -203rd St

W -204th St

W -205th St

W -206th St

W -207th St

W -208th St

W -209th St

W -210th St

W -211st St

W -212nd St

W -213rd St

W -214th St

W -215th St

W -216th St

W -217th St

W -218th St

W -219th St

W -220th St

W -221st St

W -222nd St

W -223rd St

W -224th St

W -225th St

W -226th St

W -227th St

W -228th St

W -229th St

W -230th St

W -231st St

W -232nd St

W -233rd St

W -234th St

W -235th St

W -236th St

W -237th St

W -238th St

W -239th St

W -240th St

W -241st St

W -242nd St

W -243rd St

W -244th St

W -245th St

W -246th St

W -247th St

W -248th St

W -249th St

W -250th St

W -251st St

W -252nd St

W -253rd St

W -254th St

W -255th St

W -256th St

W -257th St

W -258th St

W -259th St

W -260th St

W -261st St

W -262nd St

W -263rd St

W -264th St

W -265th St

W -266th St

W -267th St

W -268th St

W -269th St

W -270th St

W -271st St

W -272nd St

W -273rd St

W -274th St

W -275th St

W -276th St

W -277th St

W -278th St

W -279th St

W -280th St

W -281st St

W -282nd St

W -283rd St

W -284th St

W -285th St

W -286th St

W -287th St

W -288th St

W -289th St

W -290th St

W -291st St

W -292nd St

W -293rd St

W -294th St

W -295th St

W -296th St

W -297th St

W -298th St

W -299th St

W -300th St

W -301st St

W -302nd St

W -303rd St

W -304th St

W -305th St

W -306th St

W -307th St

W -308th St

W -309th St

W -310th St

W -311st St

W -312nd St

W -313rd St

W -314th St

W -315th St

W -316th St

W -317th St

W -318th St

W -319th St

W -320th St

W -321st St

W -322nd St

W -323rd St

W -324th St

W -325th St

W -326th St

W -327th St

W -328th St

W -329th St

W -330th St

W -331st St

W -332nd St

W -333rd St

W -334th St

W -335th St

W -336th St

W -337th St

W -338th St

W -339th St

W -340th St

W -341st St

W -342nd St

W -343rd St

W -344th St

W -345th St

W -346th St

W -347th St

W -348th St

W -349th St

W -350th St

W -351st St

W -352nd St

W -353rd St

W -354th St

W -355th St

W -356th St

W -357th St

W -358th St

W -359th St

W -360th St

W -361st St

W -362nd St

W -363rd St

W -364th St

W -365th St

W -366th St

W -367th St

W -368th St

W -369th St

W -370th St

W -371st St

W -372nd St

W -373rd St

W -374th St

W -375th St

W -376th St

W -377th St

W -378th St

W -379th St

W -380th St

W -381st St

W -382nd St

W -383rd St

W -384th St

W -385th St

W -386th St

W -387th St

W -388th St

W -389th St

W -390th St

W -391st St

W -392nd St

W -393rd St

W -394th St

W -395th St

W -396th St

W -397th St

W -398th St

W -399th St

W -400th St

W -401st St

W -402nd St

W -403rd St

W -404th St

W -405th St

W -406th St

W -407th St

W -408th St

W -409th St

W -410th St

W -411st St

W -412nd St

W -413rd St

W -414th St

W -415th St

W -416th St

W -417th St

W -418th St

W -419th St

W -420th St

W -421st St

W -422nd St

W -423rd St

W -424th St

W -425th St

W -426th St

W -427th St

W -428th St

W -429th St

W -430th St

W -431st St

W -432nd St

W -433rd St

W -434th St

W -435th St

W -436th St

W -437th St

W -438th St

W -439th St

W -440th St

W -441st St

W -442nd St

W -443rd St

W -444th St

W -445th St

W -446th St

W -447th St

W -448th St

W -449th St

W -450th St

W -451st St

W -452nd St

W -453rd St

W -454th St

W -455th St

W -456th St

W -457th St

W -458th St

W -459th St

W -460th St

W -461st St

W -462nd St

W -463rd St

W -464th St

W -465th St

W -466th St

W -467th St

W -468th St

W -469th St

W -470th St

W -471st St

W -472nd St

W -473rd St

W -474th St

W -475th St

W -476th St

W -477th St

W -478th St

W -479th St

W -480th St

W -481st St

W -482nd St

W -483rd St

W -484th St

W -485th St

W -486th St

W -487th St

W -488th St

W -489th St

W -490th St

W -491st St

W -492nd St

W -493rd St

W -494th St

W -495th St

W -496th St

W -497th St

W -498th St

W -499th St

W -500th St

W -501st St

W -502nd St

W -503rd St

W -504th St

W -505th St

W -506th St

W -507th St

W -508th St

W -509th St

W -510th St

W -511st St

W -512nd St

W -513rd St

W -514th St

W -515th St

W -516th St

W -517th St

W -518th St

W -519th St

W -520th St

W -521st St

W -522nd St

W -523rd St

W -524th St

W -525th St

W -526th St

W -527th St

W -528th St

W -529th St

W -530th St

W -531st St

W -532nd St

W -533rd St

W -534th St

W -535th St

W -536th St

W -537th St

W -538th St

W -539th St

W -540th St

W -541st St

W -542nd St

W -543rd St

W -544th St

W -545th St

W -546th St

W -547th St

W -548th St

W -549th St

W -550th St

W -551st St

W -552nd St

W -553rd St

W -554th St

W -555th St

W -556th St

W -557th St

W -558th St

W -559th St

W -560th St

W -561st St

W -562nd St

W -563rd St

W -564th St

W -565th St

W -566th St

W -567th St

W -568th St

W -569th St

W -570th St

W -571st St

W -572nd St

W -573rd St

W -574th St

W -575th St

W -576th St

W -577th St

W -578th St

W -579th St

W -580th St

W -581st St

W -582nd St

W -583rd St

W -584th St

W -585th St

W -586th St

W -587th St

W -588th St

W -589th St

W -590th St

W -591st St

W -592nd St

W -593rd St

W -594th St

W -595th St

W -596th St

W -597th St

W -598th St

W -599th St

W -600th St

W -601st St

W -602nd St

W -603rd St

W -604th St

W -605th St

W -606th St

W -607th St

W -608th St

W -609th St

W -610th St

W -611st St

W -612nd St

W -613rd St

W -614th St

W -615th St

W -616th St

W -617th St

W -618th St

W -619th St

W -620th St

W -621st St

W -622nd St

W -623rd St

W -624th St

W -625th St

W -626th St

W -627th St

W -628th St

W -629th St

W -630th St

W -631st St

W -632nd St

W -633rd St

W -634th St

W -635th St

W -636th St

W -637th St

W -638th St

W -639th St

W -640th St

W -641st St

W -642nd St

W -643rd St

W -644th St

W -645th St

W -646th St

W -647th St

W -648th St

W -649th St

W -650th St

W -651st St

W -652nd St

W -653rd St

W -654th St

W -655th St

W -656th St

W -657th St

W -658th St

W -659th St

W -660th St

W -661st St

W -662nd St

W -663rd St

W -664th St

W -665th St

W -666th St

W -667th St

W -668th St

W -669th St

W -670th St

W -671st St

W -672nd St

W -673rd St

W -674th St

W -675th St

W -676th St

W -677th St

W -678th St

W -679th St

W -680th St

W -681st St

W -682nd St

W -683rd St

W -684th St

W -685th St

W -686th St

W -687th St

W -688th St

W -689th St

W -690th St

W -691st St

W -692nd St

W -693rd St

W -694th St

W -695th St

W -696th St

W -697th St

W -698th St

W -699th St

W -700th St

W -701st St

W -702nd St

W -703rd St

W -704th St

W -705th St

W -706th St

W -707th St

W -708th St

W -709th St

W -710th St

W -711st St

W -712nd St

W -713rd St

W -714th St

W -715th St

W -716th St

W -717th St

W -718th St

W -719th St

W -720th St

W -721st St

W -722nd St

W -723rd St

W -724th St

W -725th St

W -726th St

W -727th St

W -728th St

W -729th St

W -730th St

W -731st St

W -732nd St

W -733rd St

W -734th St

W -735th St

W -736th St

W -737th St

W -738th St

W -739th St

W -740th St

W -741st St

W -742nd St

W -743rd St

W -744th St

W -745th St

W -746th St

W -747th St

W -748th St

W -749th St

W -750th St

W -751st St

W -752nd St

W -753rd St

W -754th St

W -755th St

W -756th St

W -757th St

W -758th St

W -759th St

W -760th St

W -761st St

W -762nd St

W -763rd St

W -764th St

W -765th St

W -766th St

W -767th St

W -768th St

W -769th St

W -770th St

W -771st St

W -772nd St

W -773rd St

W -774th St

W -775th St

W -776th St

W -777th St

W -778th St

W -779th St

W -780th St

W -781st St

W -782nd St

W -783rd St

W -784th St

W -785th St

W -786th St

W -787th St

W -788th St

W -789th St

W -790th St

W -791st St

W -792nd St

W -793rd St

W -794th St

W -795th St

W -796th St

W -797th St

W -798th St

W -799th St

W -800th St

W -801st St

W -802nd St

W -803rd St

W -804th St

W -805th St

W -806th St

W -807th St

W -808th St

W -809th St

W -810th St

W -811st St

W -812nd St

W -813rd St

W -814th St

W -815th St

W -816th St

W -817th St

W -818th St

W -819th St

W -820th St

W -821st St

W -822nd St

W -823rd St

W -824th St

W -825th St

W -826th St

W -827th St

W -828th St

W -829th St

W -830th St

W -831st St

W -832nd St

W -833rd St

W -834th St

W -835th St

W -836th St

W -837th St

W -838th St

W -839th St

W -840th St

W -841st St

W -842nd St

W -843rd St

W -844th St

W -845th St

W -846th St

W -847th St

W -848th St

W -849th St

W -850th St

W -851st St

W -852nd St

W -853rd St

W -854th St

W -855th St

W -856th St

W -857th St

W -858th St

W -859th St

W -860th St

W -861st St

W -862nd St

W -863rd St

W -864th St

W -865th St

W -866th St

W -867th St

W -868th St

W -869th St

W -870th St

W -871st St

W -872nd St

W -873rd St

W -874th St

W -875th St

W -876th St

W -877th St

W -878th St

W -879th St

W -880th St

W -881st St

W -882nd St

W -883rd St

W -884th St

W -885th St

W -886th St

W -887th St

W -888th St

W -889th St

W -890th St

W -891st St

W -892nd St

W -893rd St

W -894th St

W -895th St

W -896th St

W -897th St

W -898th St

W -899th St

W -900th St

W -901st St

W -902nd St

W -903rd St

W -904th St

W -905th St

W -906th St

W -907th St

W -908th St

W -909th St

W -910th St

W -911st St

W -912nd St

W -913rd St

W -914th St

W -915th St

W -916th St

W -917th St

W -918th St

W -919th St

W -920th St

W -921st St

W -922nd St

W -923rd St

W -924th St

W -925th St

W -926th St

W -927th St

W -928th St

W -929th St

W -930th St

W -931st St

W -932nd St

W -933rd St

W -934th St

W -935th St

W -936th St

W -937th St

W -938th St

W -939th St

W -940th St

W -941st St

W -942nd St

W -943rd St

W

Algorithm

Adding depth & normal map &
maximal Poisson-disk sampling

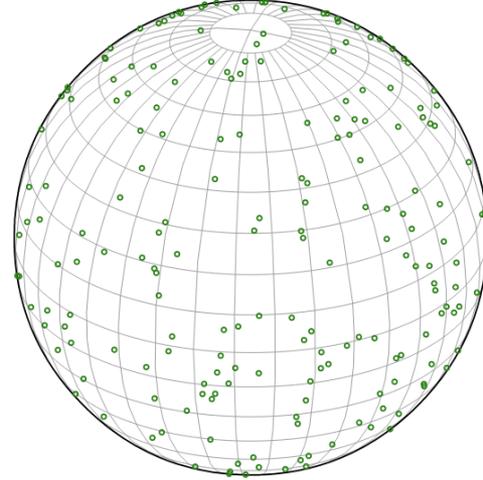
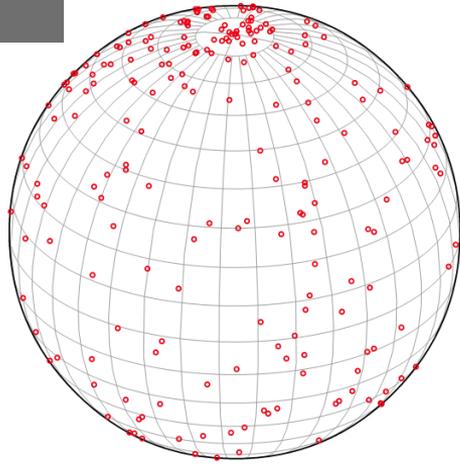


#Japanese Classic #sweets #dessert #nyc #datenight @erser_wilko
Instagram ❤️ 9 🗨️ 1 📍 NYC 2014-04-26 09:16:41
40759983416-73.982589143 (2m away) 🌟 4128' />

Social Street View enables users to see-through the nearby restaurants.

Baseline

Random Uniform Sampling



$$\varphi_i = \left(u_i - \frac{1}{2}\right)\pi, \quad \lambda_i = (2v_i - 1)\pi$$

$$x_i = \cos \varphi_i \cos \lambda_i, \quad y_i = \sin \varphi_i, \quad z_i = \cos \varphi_i \sin \lambda_i$$

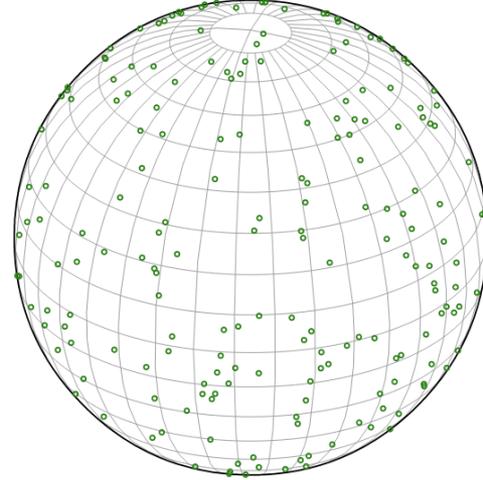
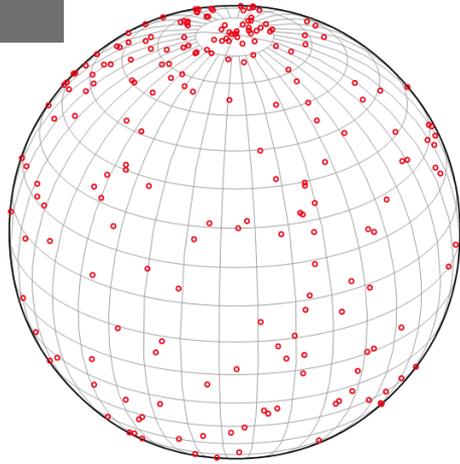
Without uniform sampling

Accumulation



Baseline

Random Uniform Sampling



$$\varphi_i = \left(u_i - \frac{1}{2}\right)\pi, \quad \lambda_i = (2v_i - 1)\pi$$

$$x_i = \cos \varphi_i \cos \lambda_i, \quad y_i = \sin \varphi_i, \quad z_i = \cos \varphi_i \sin \lambda_i$$

Not preferred
Overlays high saliency regions



Add depth map

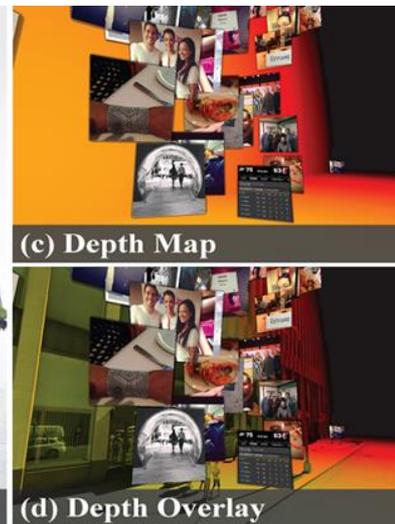
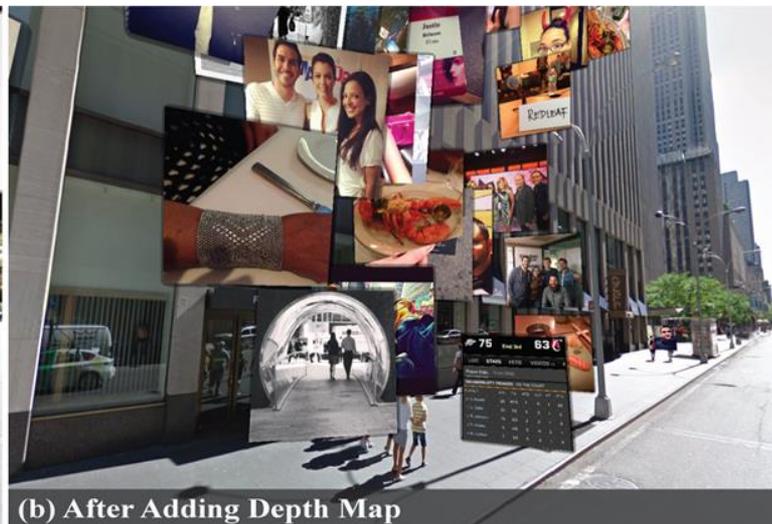
Remove sky and ground (most)

$$\Omega_s = \{q_i \mid \forall q_i \in \Omega \wedge d_i = \infty\}$$

$$\forall \tilde{p}_i \in T, D_{min} < \tilde{d}_i < D_{max}$$

Add depth map

Remove sky and ground (most)



*Can we ensure each image
aligns with the building geometries?*

Add normal map

Remove all ground, align images

$$\Omega_g = \{q_i \mid \forall q_i \in \Omega \wedge \|\mathbf{n}_i - \mathbf{n}_g\| < \delta\}$$



(a) Before Applying Normal Maps



(b) After Applying Normal Maps



(c) Depth Map



(d) Normal Map



(e) Depth+Normal Map

Can we reduce visual clutter and occlusion?

Maximal Poisson-disk Sampling

Gamito et al. Remove visual clutter and occlusion

$$\forall \tilde{p}_i \in \tilde{P}, \tilde{P} \subseteq T, \forall S \subseteq \Omega : Pr(\tilde{q}_i \in S) = \int_S di \quad (10)$$

$$\forall \tilde{p}_i, \tilde{p}_j \in \tilde{P}, \tilde{p}_i \neq \tilde{p}_j : \|p_i - p_j\| \geq r \quad (11)$$

$$S(X) = \{\tilde{p}_j \in T : \|\tilde{p}_i - \tilde{p}_j\| \geq r, \tilde{p}_i \in \tilde{P}\} : S(X) = \emptyset \quad (12)$$

Dart-throwing Algorithm

PixelPie by Ip et al. using vertex and fragment shaders



*image courtesy:
PeterPan22 @Wikimedia*

Pixel-Pie Algorithm

Remove when occlusion occurs

Algorithm 1 Maximal Poisson-disk sampling by dart-throwing

Input: The minimum distance r between sampled points

Output: A set \tilde{P} of points which satisfy equation (10)-(12)

- 1: Set $\tilde{P} \leftarrow \emptyset$, empty region $\tilde{R} \leftarrow T$
 - 2: **repeat**
 - 3: Generate some random points $\tilde{P}' \subseteq R$ by rasterizing them as circular disks into a depth map in vertex shader.
 - 4: Remove any point $\tilde{p} \in \tilde{P}'$ whose corresponding point \tilde{q} violates $\tilde{q} \in \Omega_g \vee D_{min} < \tilde{d} < D_{max}$
 - 5: Identify and remove the occluded disks from \tilde{P} by reading the depth map in the shader.
 - 6: $\tilde{P} \leftarrow \tilde{P} \cup \tilde{P}'$
 - 7: Update the empty region R in the fragment shader.
 - 8: **until** $R \leftarrow \emptyset$
-

Project Social Media Pictures

By Maximal Poisson-disc Sampling

ALGORITHM 1: Social Media Layout using Poisson-disk Samples

Input: N sorted social media images $\hat{S} = \{s_i \mid i = 1 \dots N\}$, acquired from SSV servers.

Output: A set of image planes to display social media: $I = I_1 \dots I_M, M \leq N$.

Generate the set of candidate sample points $\tilde{\mathbf{P}}$ by the PixelPie algorithm;

Sort points in $\tilde{\mathbf{P}}$ in descending order according to their corresponding values in the depth map D so that the closest sample point is laid out first;

Set $I \leftarrow \emptyset$;

for $i \leftarrow 1 \dots \min(N, |\tilde{\mathbf{P}}|)$ **do**

Place I_i with texture from $s_i \in \hat{S}$ at the projected position $\tilde{\mathbf{q}}_i \leftarrow \mathcal{P}(\tilde{\mathbf{p}}_i)$;

Rescale I_i according to the corresponding depth value: $\tau_i \leftarrow \tau/d_i$ for perspective visual effects;

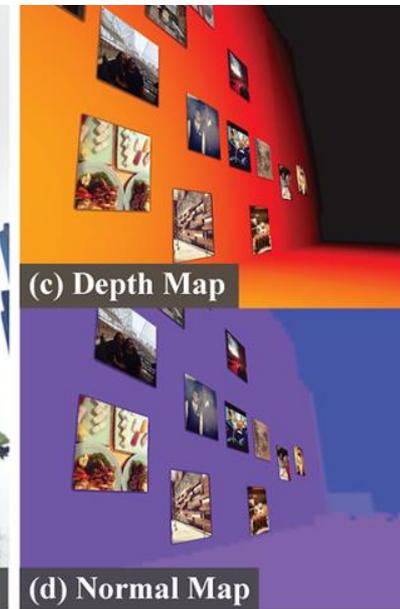
Rotate I_i so that it is perpendicular to the normal vector $\mathbf{n}_i \leftarrow \mathbf{N}(u_i, v_i)$;

Add I_i to the result set: $I \leftarrow I \cup I_i$;

end

Sampling Comparison

Remove visual clutter and occlusion



Scenic Landscapes

Using orientation of the road

ALGORITHM 2: Social Media Layout using Road Orientations

Input: $|O|$ road orientations with $o_i \in [0, 2\pi]$. K social media to be placed for each orientation. Typically, $|O| = 2$ for a road with two orientations.

Output: A set of image planes to display social media:

$$I = I_1 \dots I_M, M \geq K \cdot |O|.$$

Set $I \leftarrow \emptyset$;

for $i \leftarrow 1 \dots |O|$ **do**

Set the position $\mathbf{q}_i \leftarrow (KR \cos o_i, h, KR \sin o_i)$ at height h and radius R ;

(Optional based on user's preference) Add a frontal image plane to I at \mathbf{q}_i ;

Set the translation $\mathbf{t} \leftarrow (T \cos(o_i + \frac{\pi}{2}), 0, T \sin(o_i + \frac{\pi}{2}))$ with constant T ;

for $k \leftarrow 1 \dots K$ **do**

Set $\tilde{\mathbf{q}} \leftarrow (kR \cos o_i, h, kR \sin o_i)$;

Add a left side image plane to I at position $\mathbf{q}' \leftarrow \tilde{\mathbf{q}} + \mathbf{t}$;

Add a right side image plane to I at position $\mathbf{q}' \leftarrow \tilde{\mathbf{q}} - \mathbf{t}$;

end

end

Scenic Landscapes

Using orientation of the road



Scenic Landscapes

Using orientation of the road



Algorithm

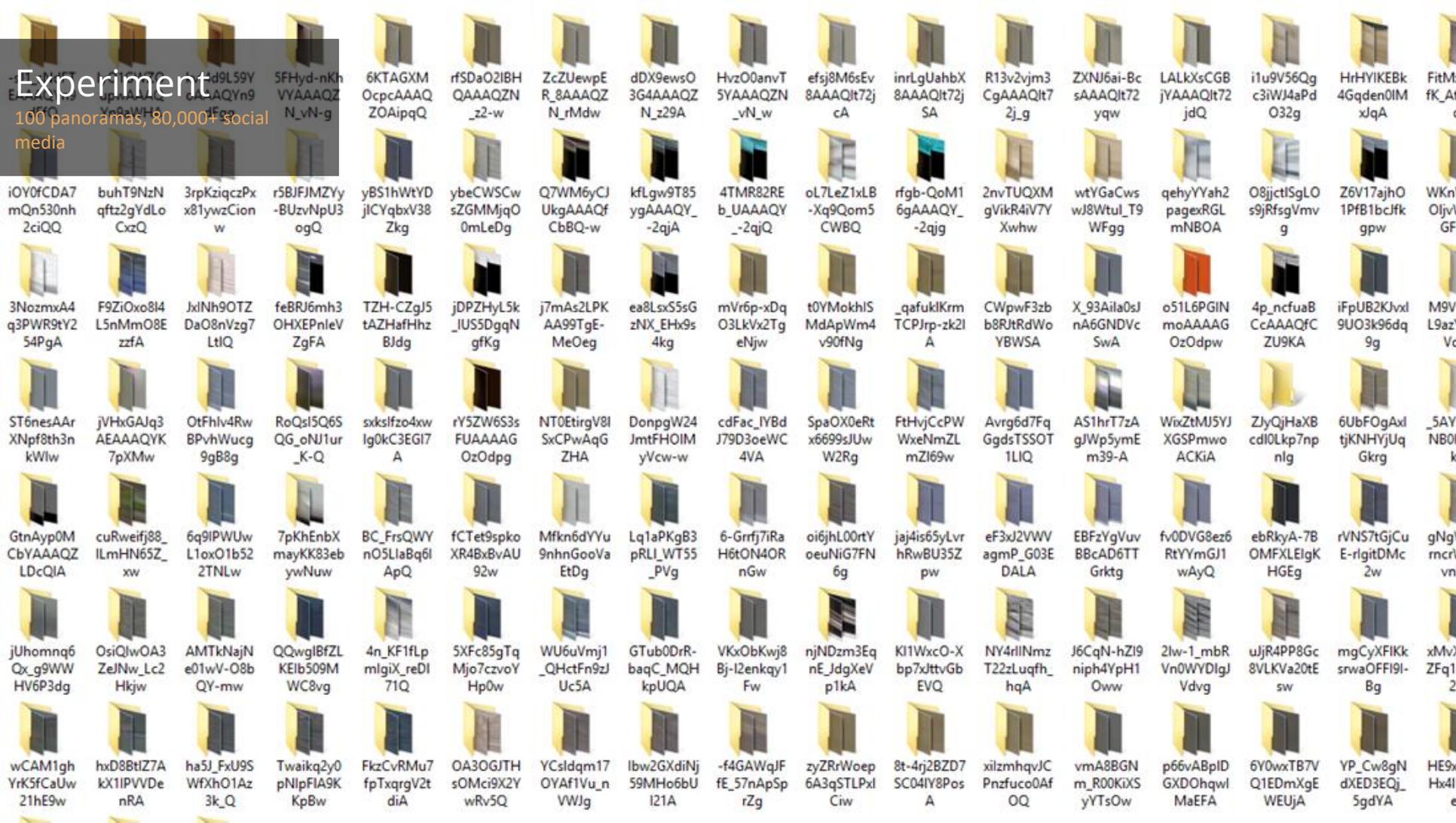
Adding depth & normal map &
maximal Poisson-disk sampling



In addition, our system allows users to walk around and explore live social media streams.

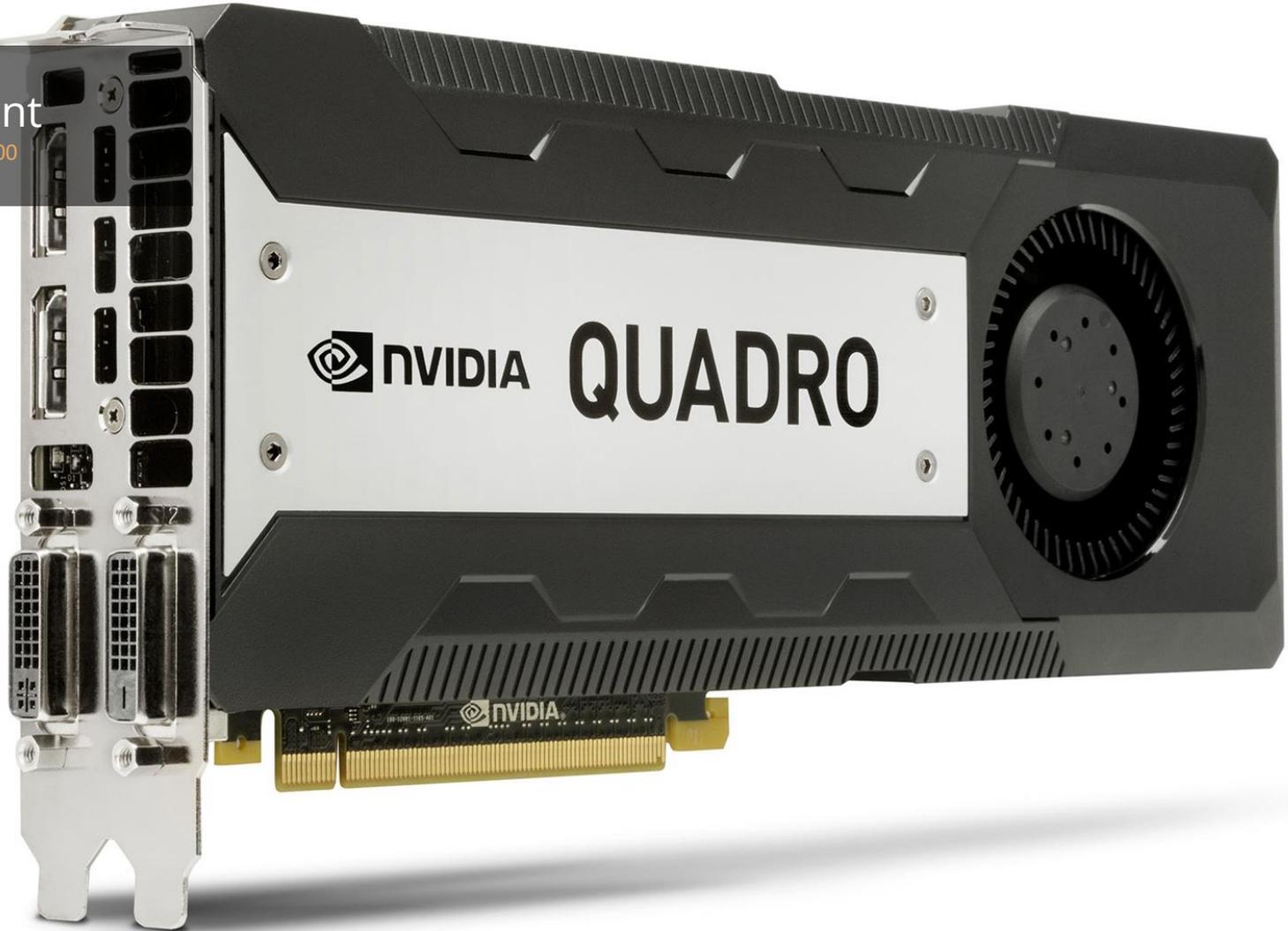
Experiment

100 panoramas, 80,000+ social media



Experiment

Nvidia Quadro K6000



Experiment

Conditions of panoramic images

Immersive high-resolution screens



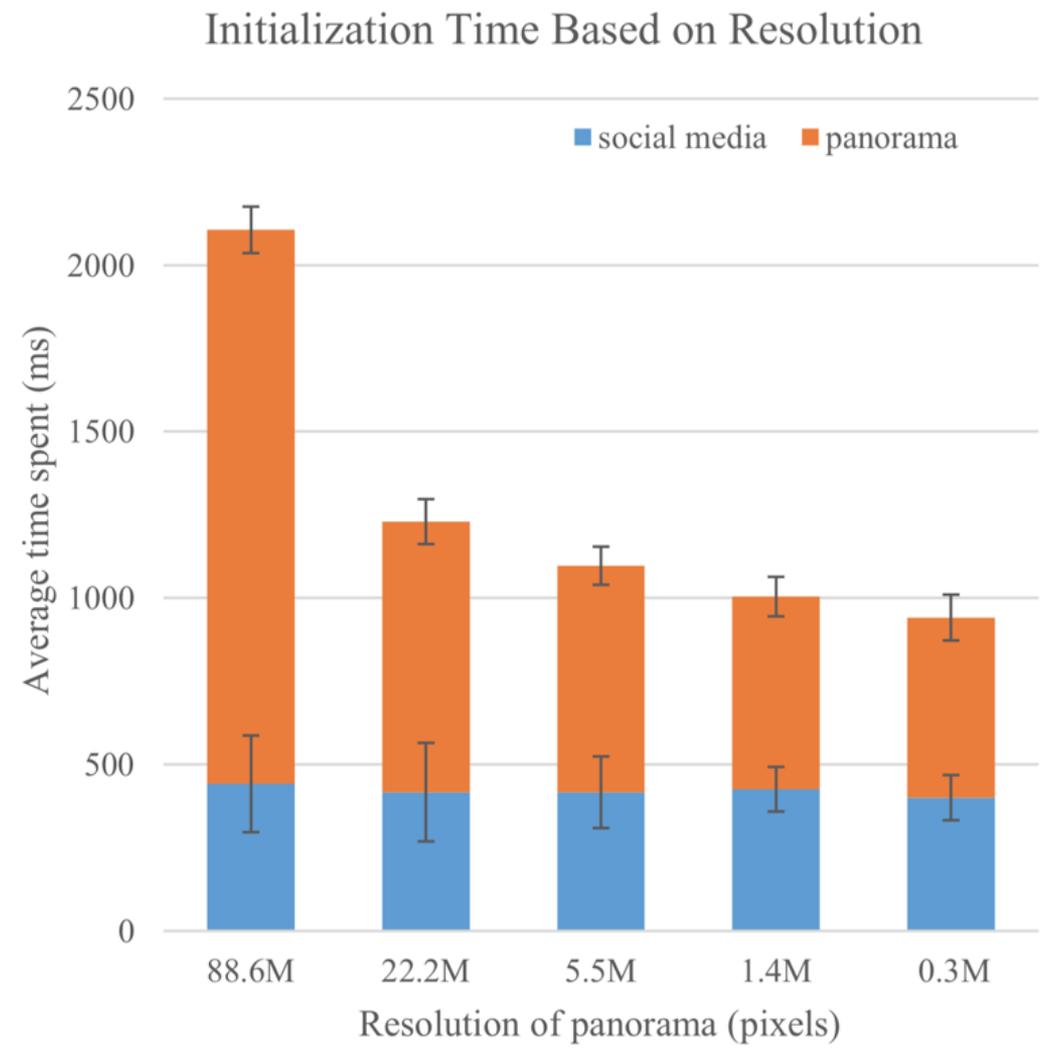
Pixels	Resolution	Number of tiles	File size
88.6M	13312×6656	26×13	$\sim 5M$
22.2M	6656×3328	13×7	$\sim 2M$
5.5M	3328×1664	7×4	$\sim 800K$
1.4M	1664×832	4×2	$\sim 300K$
0.3M	832×416	2×1	$\sim 90K$

Common Consumer-level Displays

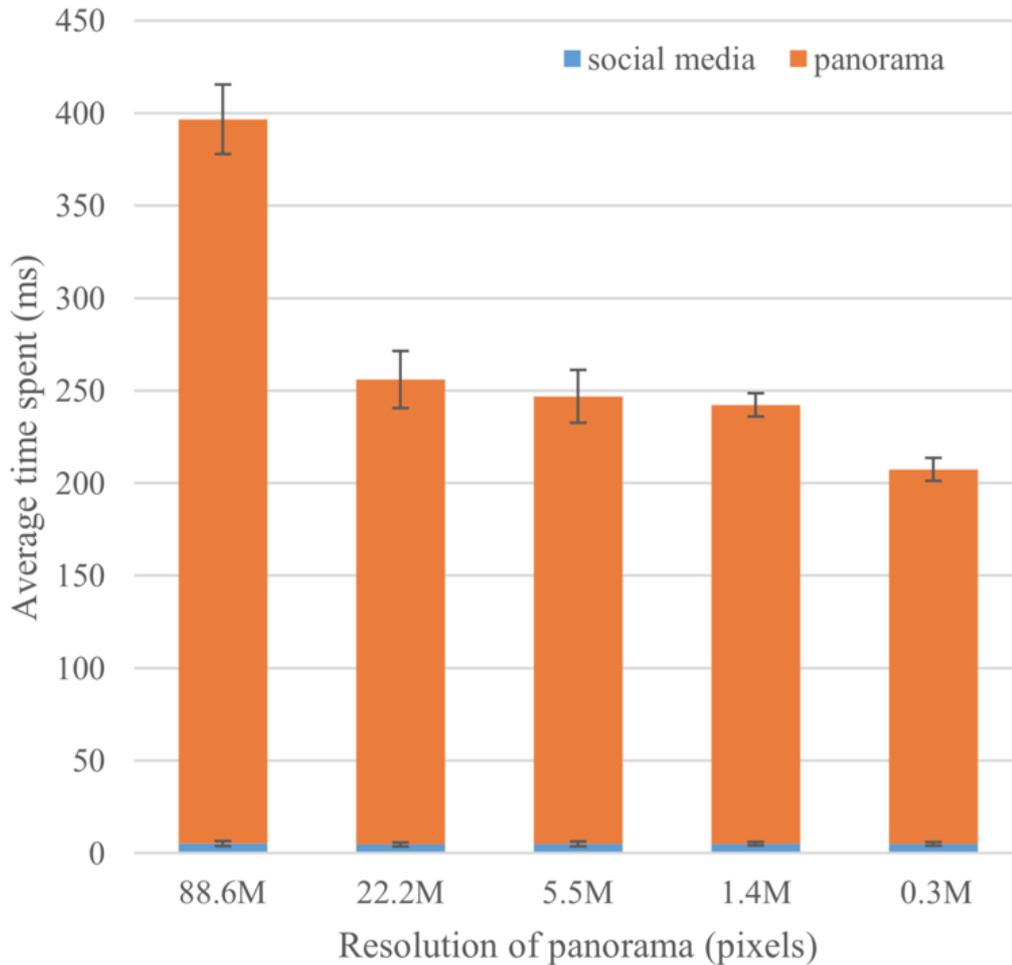


Initialization Time

Panorama takes a while to load



Initialization Time After Prefetching

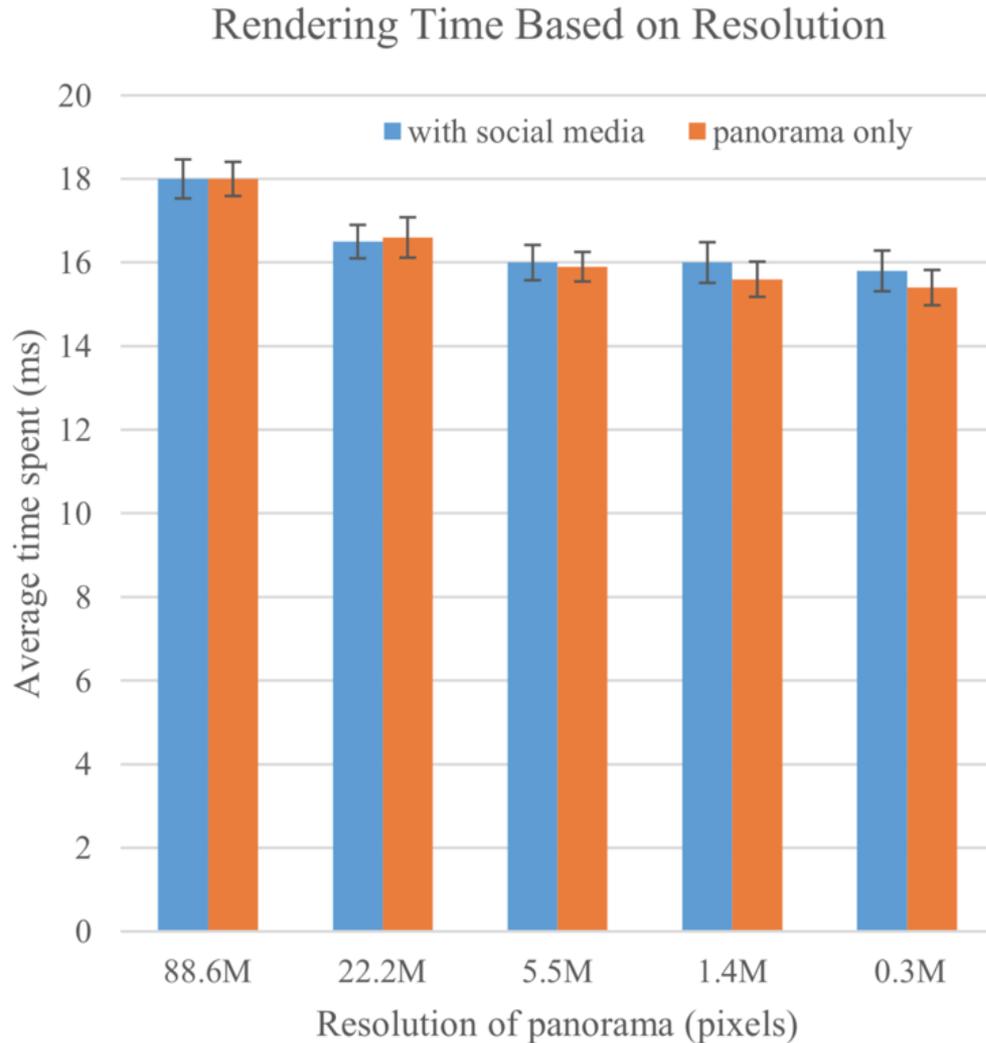


After Prefetching

¾ - ¾ time reduced

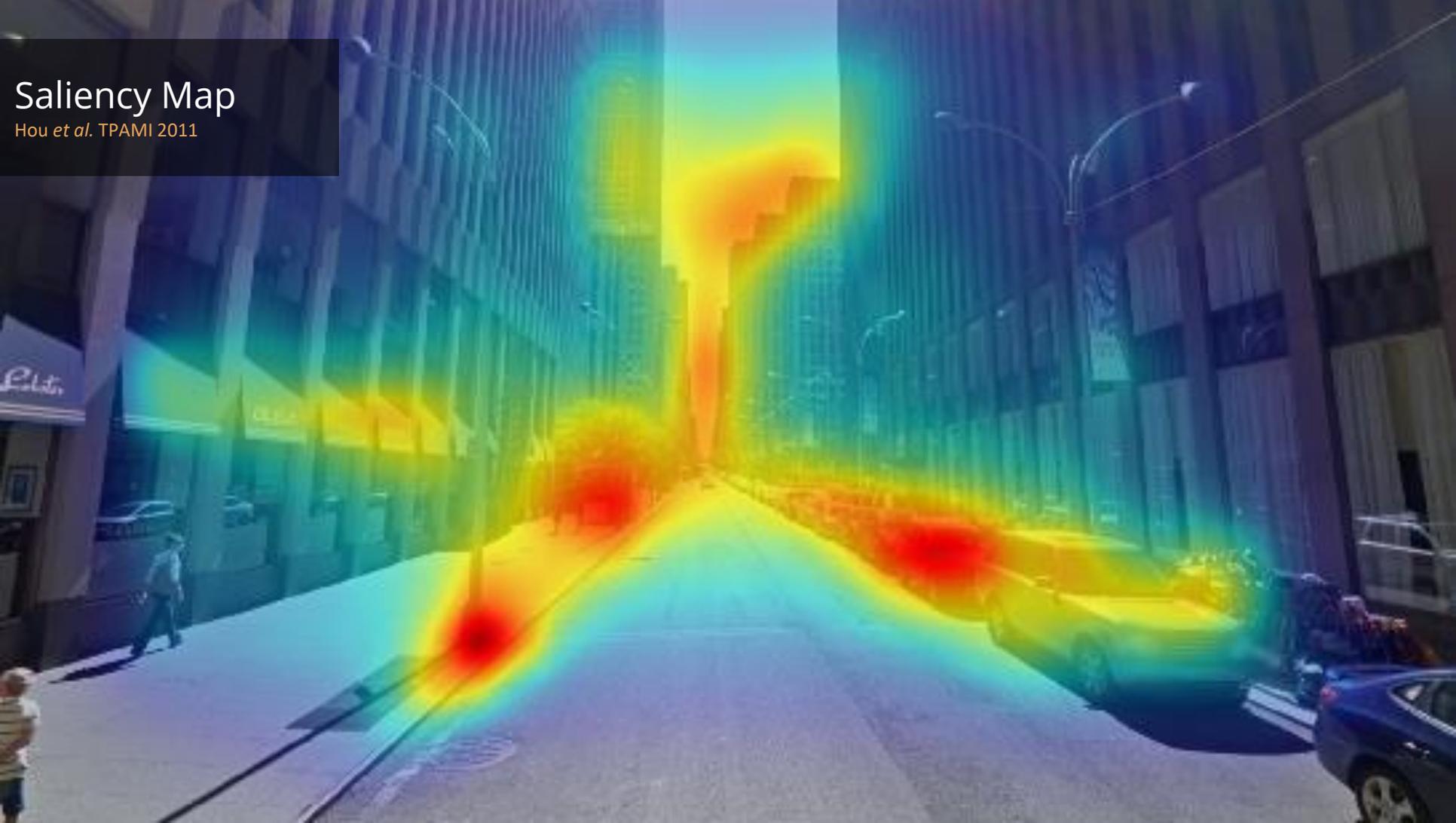
Rendering Time

Almost 58~60 FPS



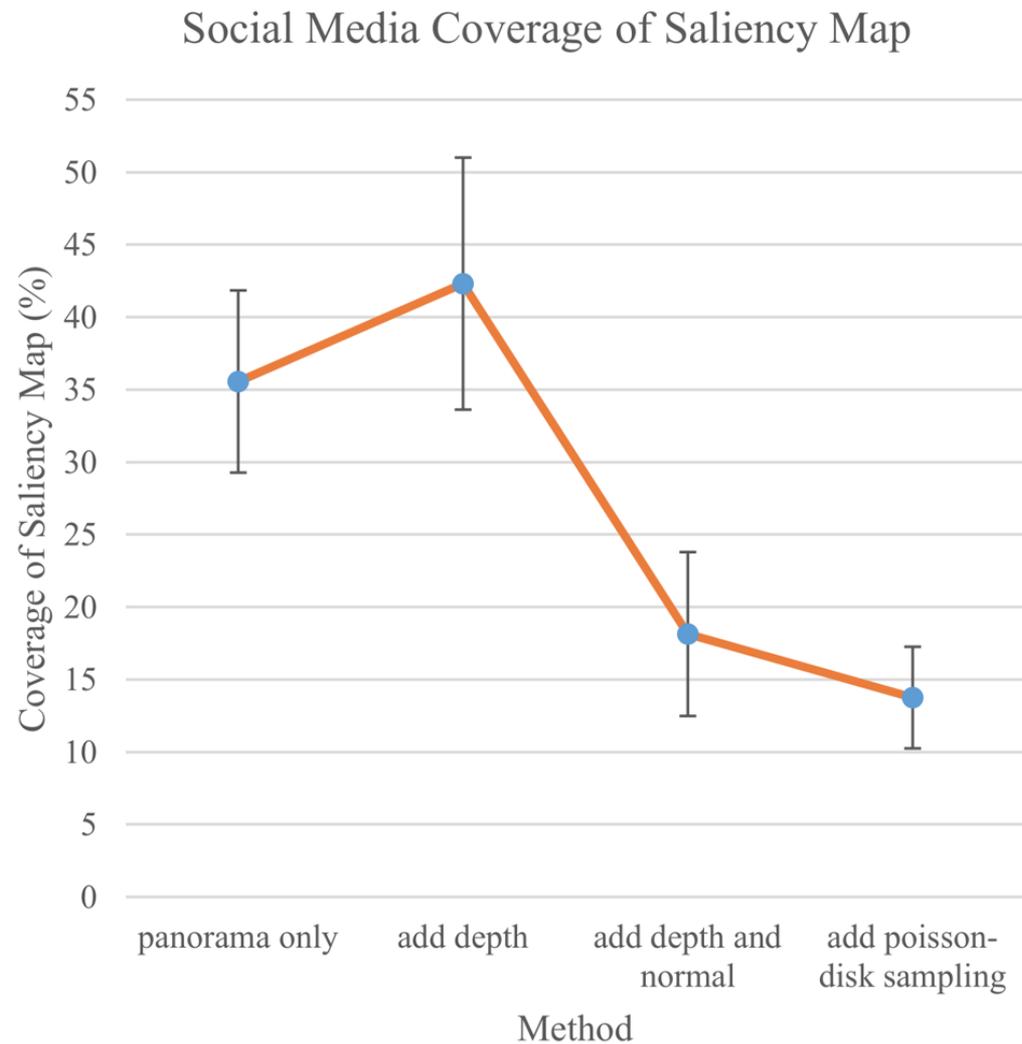
Saliency Map

Hou *et al.* TPAMI 2011



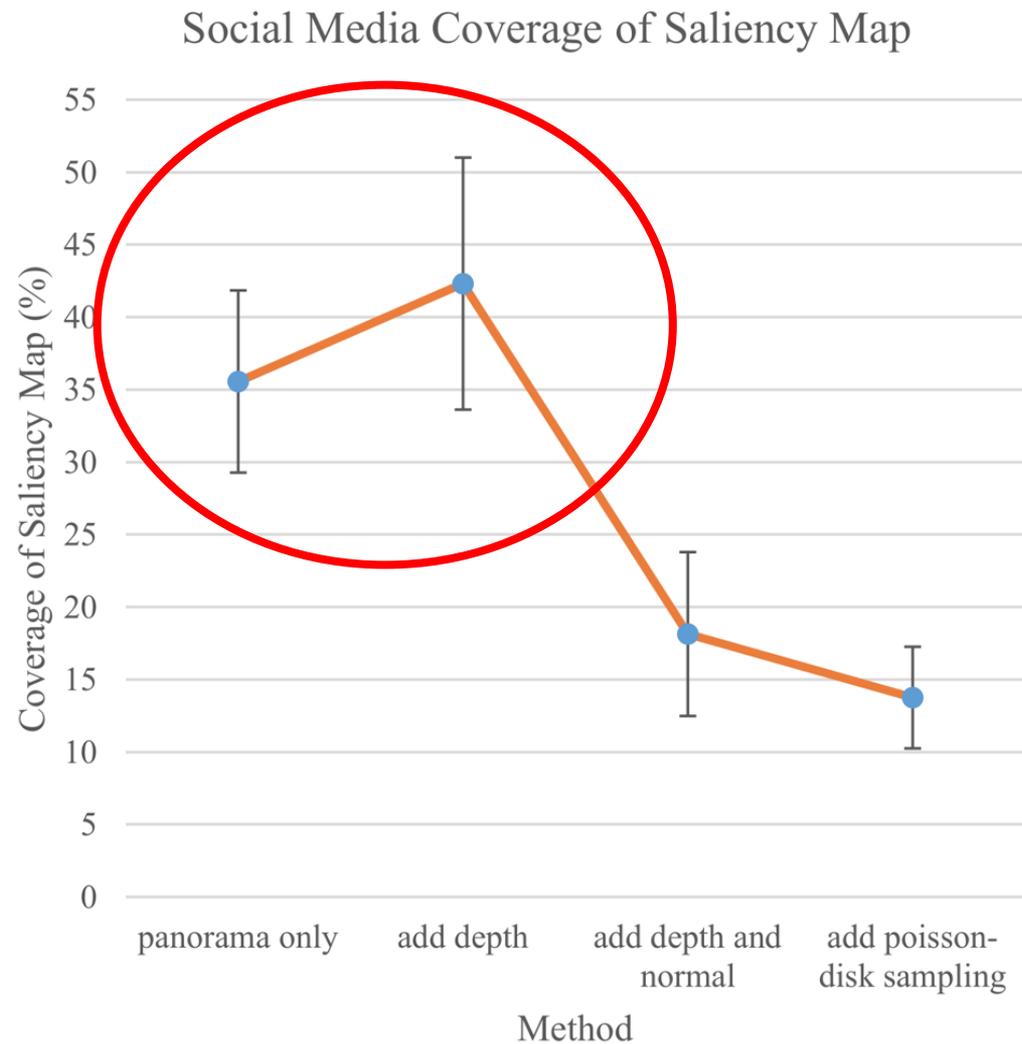
Social Media Coverage

100 panoramas for each algorithm



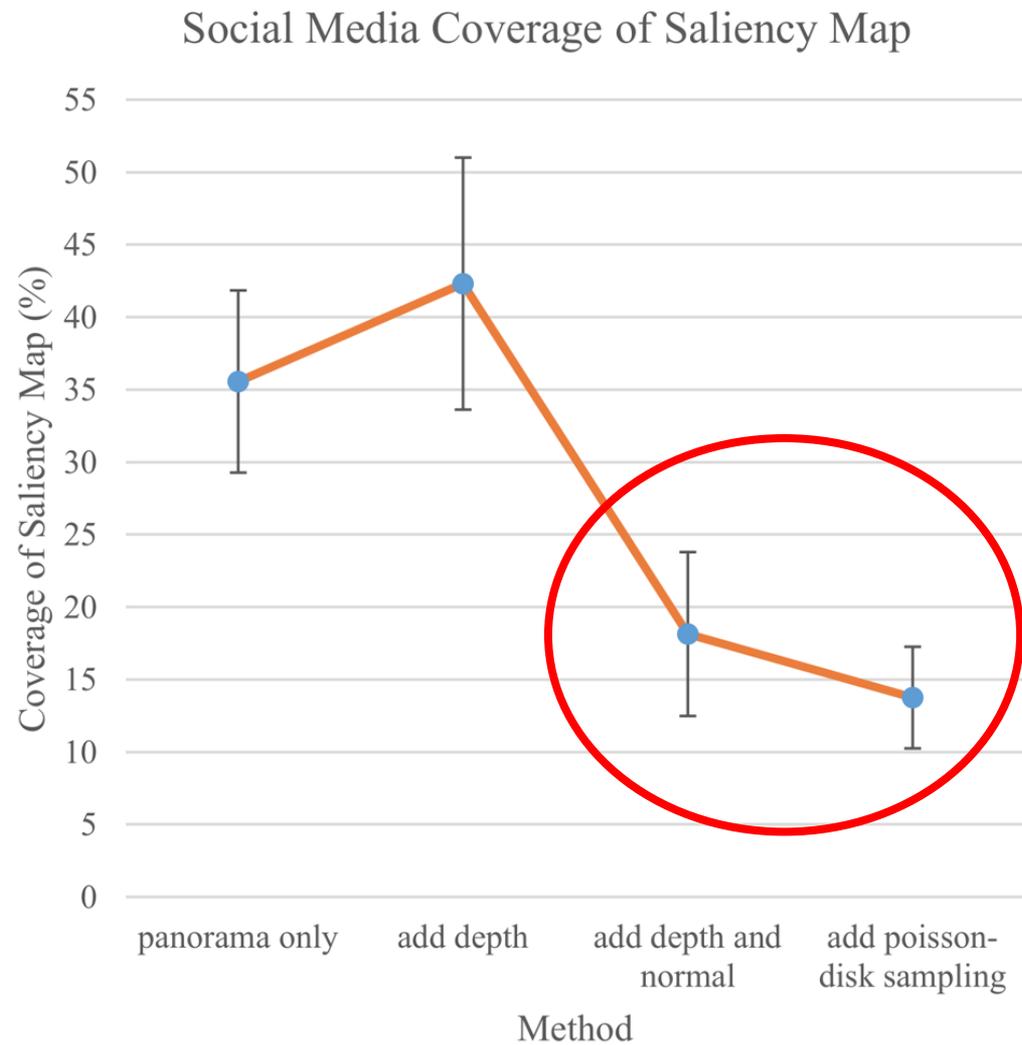
Social Media Coverage

100 panoramas for each algorithm



Social Media Coverage

100 panoramas for each algorithm



Potential Applications?



“

Stuck in traffic on our way to
Cabo with this awesome view

#roadtrip #cabo #view #mexico

”

Daniela on *Instagram*
July 12, 2014



Application

Immersive story telling



However, we can hardly enjoy the view given the small posted image.

Business Advertising

Museum, restaurant, real-estate ...



“

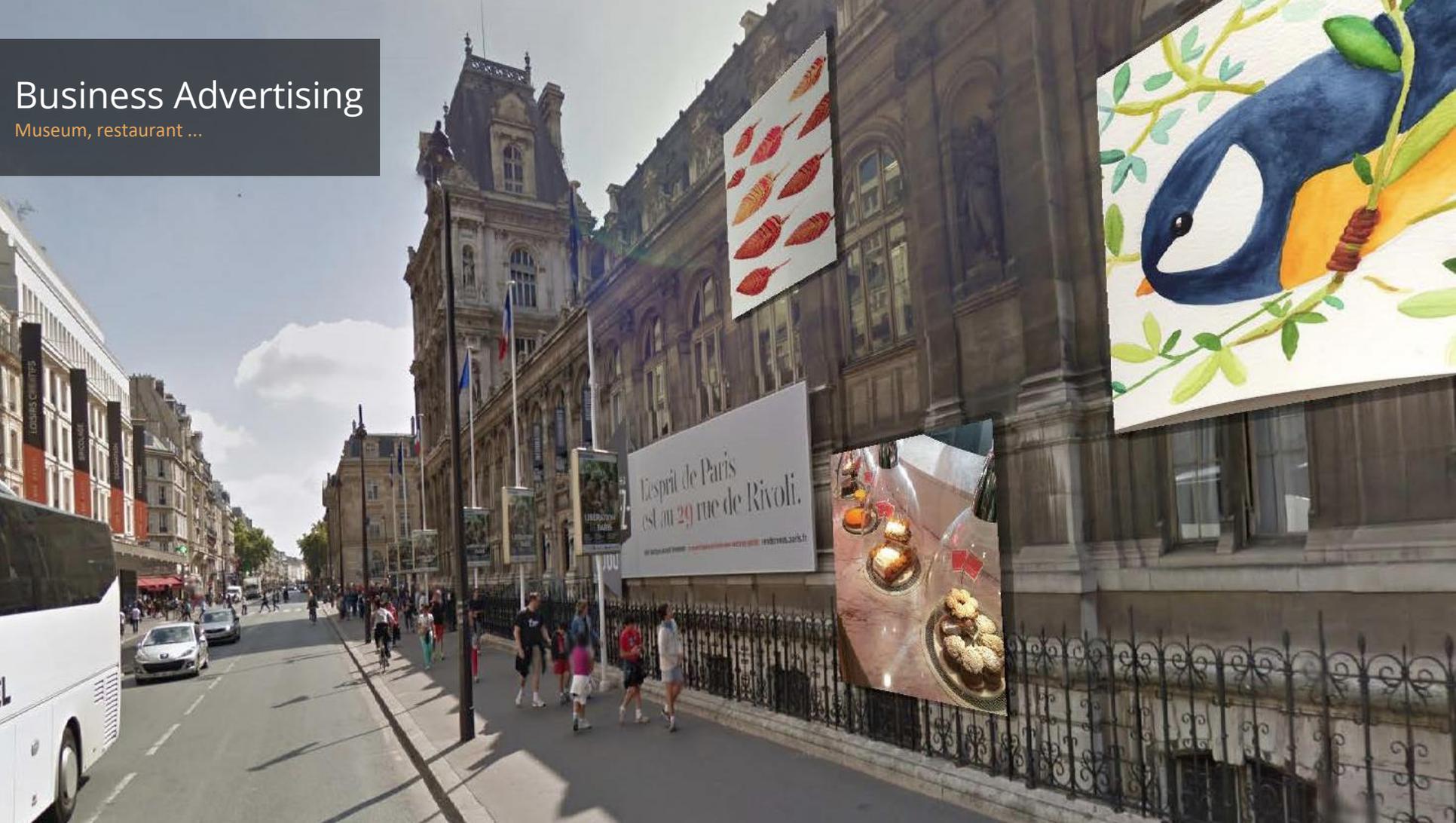
... dinner started off with
amazing oysters paired with my
favorite Ruinart blanc de
blancs champagne

”

By frankiextah on
Instagram

Business Advertising

Museum, restaurant ...



Learning Culture

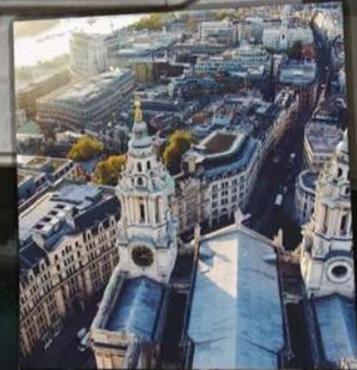
Taierzhuang, Chinese Spring Festivals



Crowd-sourced Tourism



Crowd-sourced Tourism





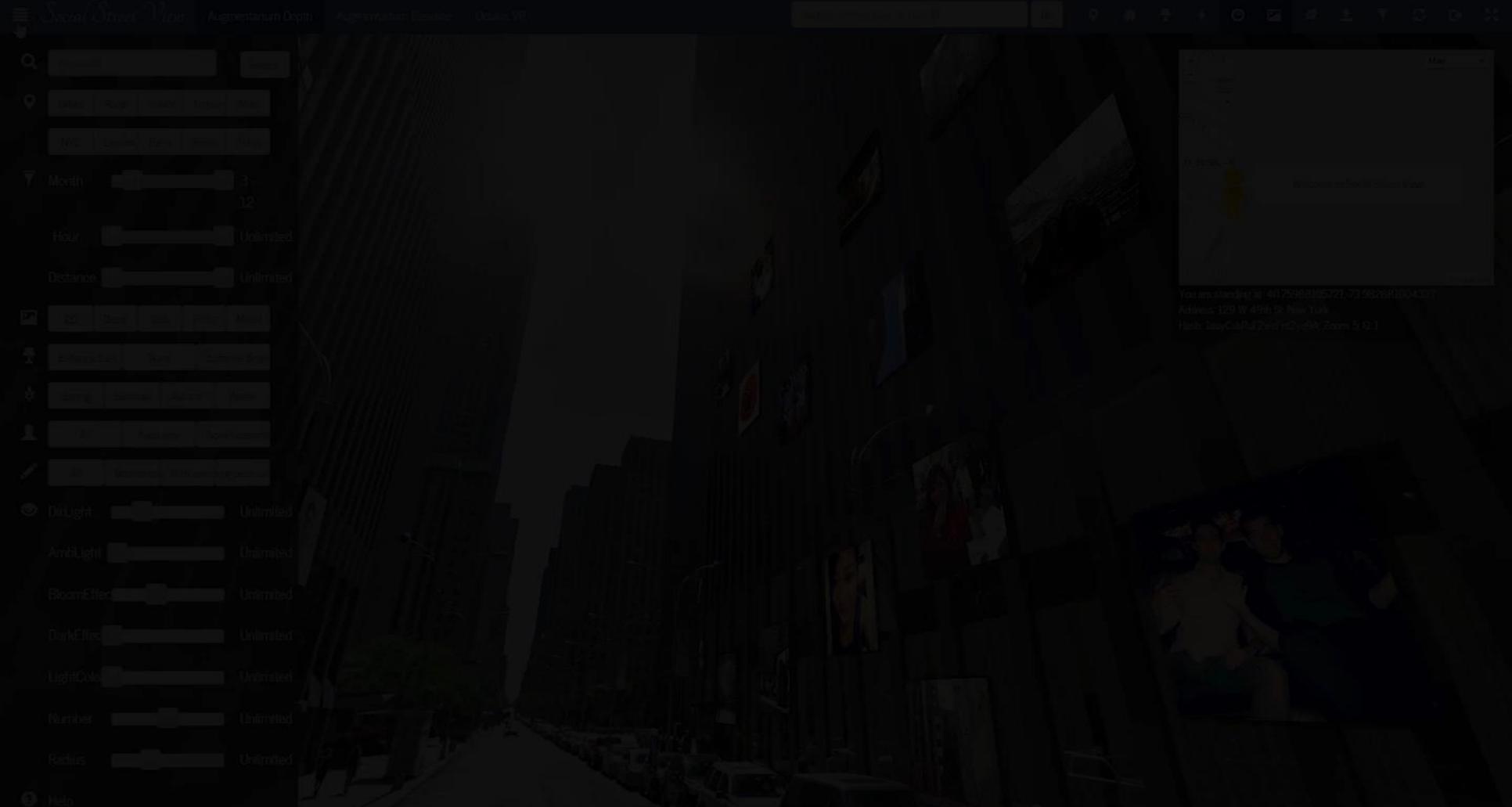
What if ...

Temporal information is used for filtering and rendering?

Visual Enhancement

Spatio-Temporal Augmented Reality





Users can use temporal or distance filters to narrow down their interested queries.

Seasonal Effect

Winter



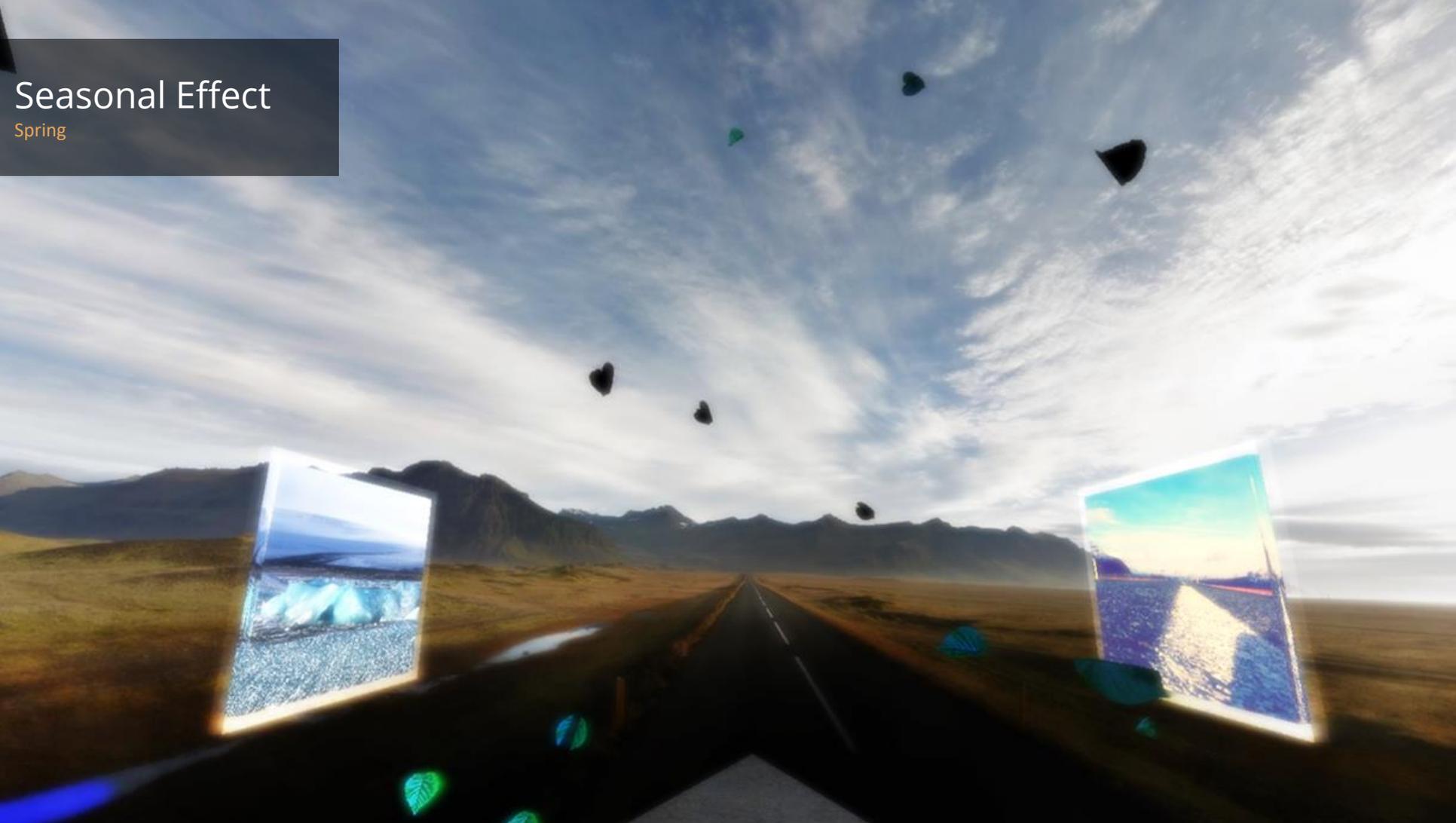
Seasonal Effect

Winter



Seasonal Effect

Spring



Seasonal Effect

Autumn in Paris



Failure Cases

A Square in London



Future Work

Digital City by 3D Reconstruction, Depth Fusion and Augmented Reality



image courtesy:
wallpapervortex.com

Future Work

Incorporate with Open3D in Web3D 2016?

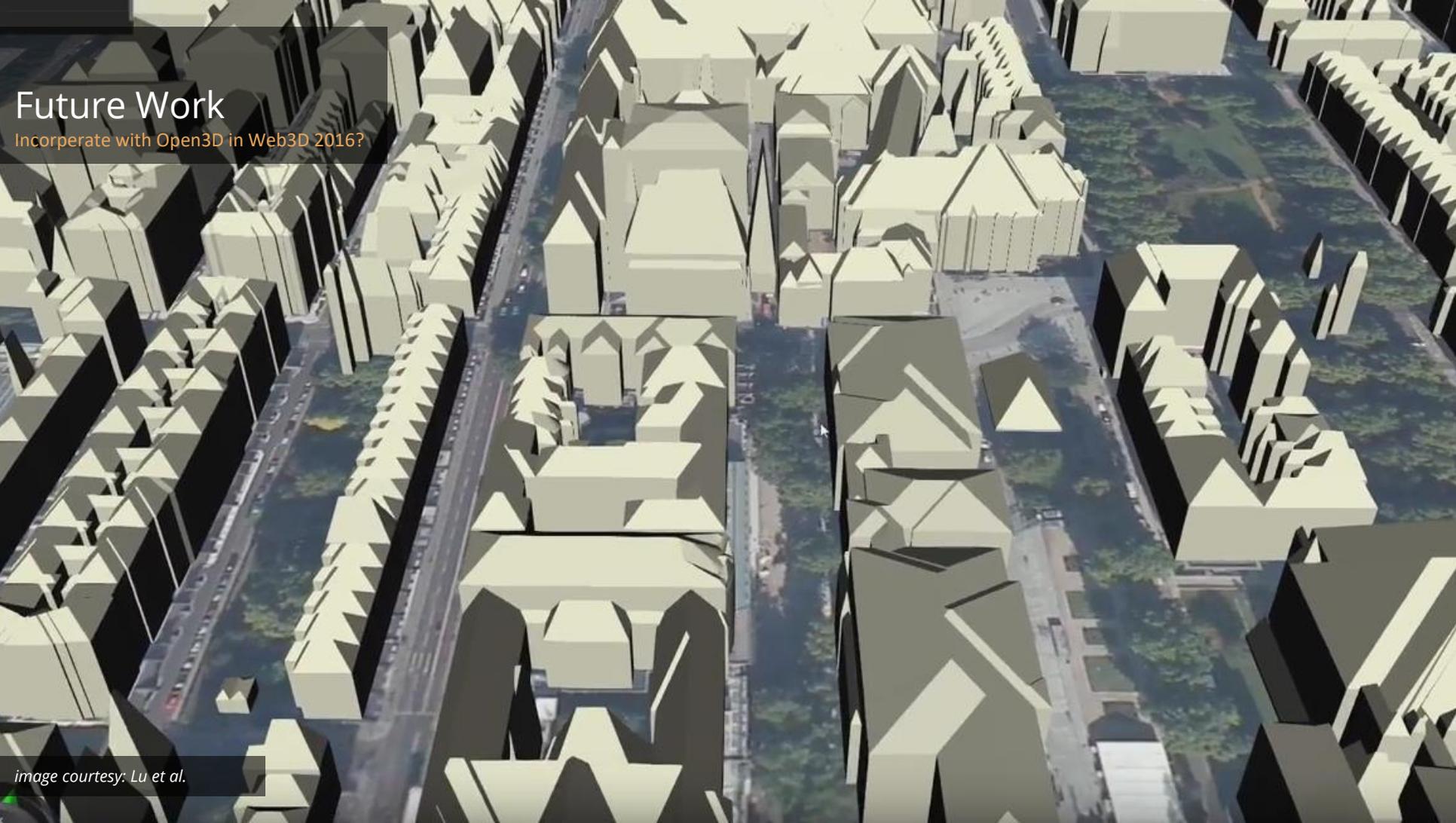
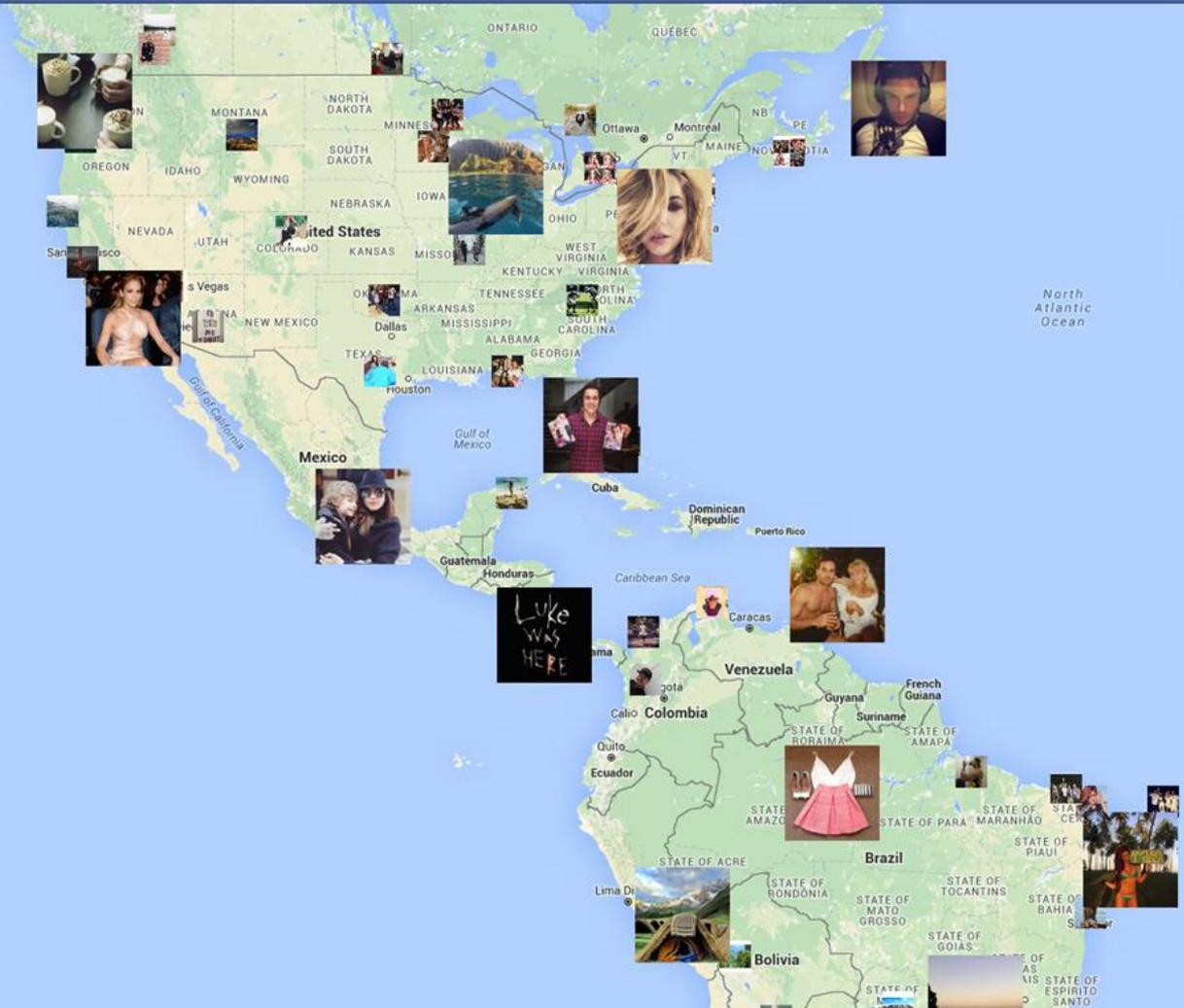


image courtesy: Lu et al.

Acknowledgement

- Visualization Style
- Help
- About



Acknowledgement

Augmentarium Lab | GVIL | UMIACS



Acknowledgement

NSF | Nvidia | MPower | UMIACS



UNIVERSITY OF
MARYLAND

UMIACS

University of Maryland
Institute for Advanced
Computer Studies



UNIVERSITY OF MARYLAND
MPOWERING THE STATE



COMPUTER SCIENCE
UNIVERSITY OF MARYLAND



Video Fields

Web3D Sunday 9:20 - 9:40 AM



Surveillance Video Streams

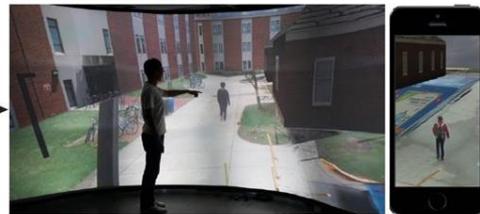


Camera World Matrix Calibration



Static 3D Models and Satellite Image

Video Fields
Mapping



Dynamic Virtual Environment



Automatic Segmentation and View-dependent Rendering

Social Street View

www.SocialStreetView.com

Thank you! Any questions or comments are welcome!

Ruofei Du and Amitabh Varshney

Augmentarium Lab | GVIL | UMIACS

Web3D 2016

Social Street View Backup Slides

Future Work

Saliency

- Low-level image saliency for layout
- Real-time or post-processed saliency maps

Future Work

User Study

- High-resolution large-area screens
- Head-mounted displays

Future Work

3D Reconstruction and AR

- Real-time geometric fusion
- HoloLens

Future Work

3D Reconstruction and AR

- What is the (experimentally-justifiable) motivation for putting generic images into scenes?
- What is the relevance of which photos are displayed? And how this affects the worth?
- Whether it is possible to try and improve this relevance using vision techniques?

What is ignored in current social media sites?