Social Pixels: Genesis and Evaluation

Vivek K. Singh
University of California, Irvine
singhv@uci.edu

Mingyan Gao
University of California, Irvine
gaom@uci.edu

Ramesh Jain
University of California, Irvine
jain@ics.uci.edus

ABSTRACT

Huge amounts of social multimedia is being created daily by a combination of globally distributed disparate sensors, including human-sensors (e.g. tweets) and video cameras. Taken together, this represents information about multiple aspects of the evolving world. Understanding the various events, patterns and situations emerging in such data has applications in multiple domains. We develop abstractions and tools to decipher various spatio-temporal phenomena which manifest themselves across such social media data. We describe an approach for aggregating social interest of users about any particular theme from any particular location into 'social pixels'. Aggregating such pixels spatio-temporally allows creation of social versions of images and videos, which then become amenable to various media processing techniques (like segmentation, convolution) to derive semantic situation information. We define a declarative set of operators upon such data to allow for users to formulate queries to visualize, characterize, and analyze such data. Results of applying these operations over an evolving corpus of millions of Twitter and Flickr posts, to answer situation-based queries in multiple application domains are promising.

Categories and Subject Descriptors

H.4.m [**Information Systems**]: Miscellaneous

Keywords

Human-sensors, microblogs, social pixel, Situation detection, query algebra, media processing

1. INTRODUCTION

We are currently witnessing an explosive growth in the social web, where large amounts of social multimedia is being created daily by a globally distributed array of disparate sensors, including human-sensors (e.g. tweets) and video cameras. Sites like Twitter, Facebook, and Flickr are reporting millions of new user posts each day. However, the current technology trends have mostly focused on *producing* large volumes of such media. The techniques on *consuming* or *utilizing*, such social multimedia are still in infancy. We argue that the piece-meal consumption of any of such tweets, images, or songs has only limited implications. Rather, the real rewards lie in utilizing aggregated collections of such media to understand various real world spatio-temporal phenomena whose aspects are captured within such social media.

While (on the surface) appearing to be individualistic and unsynchronized, the social web is known to demonstrate critical self organizing behavior[1]. The patterns emerging often show deep interconnections with various world events [19, 23] and in effect capture the evolving world model at each instant. Hence we consider the emerging social web to be a *macroscope* wherein millions of *human-sensors* across the globe are capturing different aspects of the spatio-temporal phenomena which can be aggregated into a holistic view.

Spatio-temporal organization and processing of such social media will have applications in domains ranging from automated trip planning, and customized recommendations at individual user level, to emergency response, political campaign management, weather analysis, and geo-spatial business intelligence at a global level. This underscores the need for tools which can effectively organize and process large volume of such social media data and make explicit the semantic information from them.

We feel that multimedia research tools are especially well suited for studying such socio-media phenomena. One, they are well positioned to handle the different forms of media (audio, video, text, sensors etc.) being created on the social web. Equally importantly, they have rich array of tools for handling both spatial (like spatial organization of pixels in images) and spatio-temporal (videos) organizations of data.

In this paper we present an approach in this direction. Taking inspiration from traditional image pixels which represented aggregation of photons at a location, we define aggregation of user interest at a particular geo-location as a 'social pixel'. Combining such social pixels spatially allows us to create *e-mages* (i.e. an event data based analog of images) and combining them across space and time leads to *temporal e-mage sets*. Such a visual representation (e.g. shown in fig 1), allows for intuitive understanding by any human user. Further an image/video based representation, allows for the use rich repository of media processing algorithms (like flow patterns, segmentation, convolution) to easily derive semantically useful situation information. Such analysis would be very tedious in a text-based corpus of similar data or even as querying based approach in traditional databases where (relatively simple) media processing operators like convolution and segmentation are yet to be mapped effectively.

We realize that the real end users of such systems (e.g. a political campaign manager, or emergency response analyst) are unlikely to be experts in the *procedural* aspects of data processing. In fact, procedural methods/ languages are known to require significant training before users can employ them, and often tend to be tool and format dependent [25] [11]. Hence, we define a set of *declarative* query operators, where the user just describes her data needs. The defined spatio-temporal query operators allow users from multiple domains to interact with the social media data and ask questions on derived attributes (e.g. velocity, epicenter of the distribution) which would not be available directly out of raw data feeds.

To summarize, our main contributions in this work are:
1. Defining an *social pixel* based approach for a unified spatio-temporal representation of social multimedia data.
2. Developing a media-processing inspired approach for deriving relevant situational attributes from such data.
3. Defining a *declarative* query algebra for an end user to query such a system.

While the ideas are generic and can be applied to any social media, in this paper we focus on 'Twitter' and 'Flickr' for their easy availability, *real-time* (especially Twitter) data characteristic, large user volume, and inherent spatio-temporal-thematic nature. To verify the efficacy of the proposed approach, we demonstrate multiple applications including business intelligence, political event analytics, and seasonal characteristics analysis on this data set.

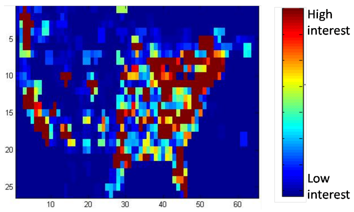

Figure 1: Example of a 'social image'. This image show interest level amongst users across mainland US in terms of number of tweets containing the term 'iphone' on 11th Jun 09.

2. DESIGN PRINCIPLES

2.1 Humans as sensors

Humans as sensors can describe aspects of a situation, which are not yet measurable by any hardware sensors. Humans can describe perceptions, emotions, impressions (for business products), counter state censors (as seen in Iran elections), be first respondents (e.g. Hudson river landing), emergency reporters (Haiti earthquake rescues), and even pass unconfirmed/unofficial information reports (rumors, merger-info, scandals). Growing importance of multimodal user contributions (e.g. twitpic, twaudio, Flickr), and location based services (e.g. Foursquare, Gowalla, Yelp) clearly highlights this. Social media networks arguably thus represent a large self organizing sensor network, more sophisticated than any other sensor network currently deployed. Hence, we intend to exploit human sensor inputs for situation awareness and control in this work.

2.2 Social pixel approach

Traditional imaging sensors employed *pixels* which represented aggregations of the various photon energies striking at any particular location. In situation sensing, we define *social pixels* to be representative aggregations of different user contributions coming a particular geo-location. For example, a large number of tweets about 'swine flu' coming from a particular geo-location, can be represented as a 'high' value at the corresponding pixel.

The abstraction of social media content into spatio-temporal 'pixels', 'images', and 'videos' has implications on multiple aspects of this work.
1. Visualization: This approach allows for intuitive visualization and hence aids situation awareness for a human user. Humans are quite used to seeing satellite image and GIS data on similar interfaces and hence can understand such data much better than any text or data-base centric representation. A large portion of human brain is dedicated to visual processing, which explains the popularity of dashboard based situation awareness tools[4].
2. Intuitive Query and mental model: The correspondence of the human mental model of spatio-temporal data with the query processing model, makes it easier for humans to pose queries, and understand the results.
3. Common representation: Such a representation allows multiple spatio-temporal data sources (e.g. Maps, weather info, demographics, geocoded twitter feeds, Flickr images) to be assimilated within the same framework.
4. Data Analysis: Such a representation allows us to exploit a rich repository of media processing algorithms which can be used to obtain relevant situational information from this data. For example, well developed processing techniques (e.g. filtering, convolution, background subtraction) exist for for obtaining relevant data characteristics in real time.

Besides these, a spatio-temporal binning is better for *individual* user privacy, and reduces run-time query processing cost.

2.3 Combining media processing with declarative query algebra

Spatio-temporal databases have their strength in being relatively easier for users to pose queries. Media processing tools, on the other hand can undertake much more sophisticated processing. Hence, in this work we want to combine the strengths of the two approaches.

3. RELATED WORK

Multiple works at the intersection of multimedia, signal processing, HCI, and Web 2.0 are becoming increasingly interested in the potential here. For example multiple ideas like 'citizen/ participatory sensing'[22], 'social signal processing', 'human computation', 'crowd-sourcing', 'collective intelligence', 'wisdom of the masses'[9], etc. have become increasingly popular recently. We contribute to such efforts by providing useful media processing based analytic tools.

There have been some attempts recently at combining social media content across users for meaningful applications. GIS based works like [8] (using a global network of webcams to detect weather patterns), Photo-synth based combination of Flickr images[24], Twitter based combination of microblogs for analyzing presidential debate[21], Swine-flu monitoring [23], and earthquake monitoring [19] are excellent examples of these. There has also been significant work recently on analyzing social networks and the blogosphere. Twitris[13] allows users to observe trending (popular) topics on Twitter, and collect the related information across media sources. Blogscope [3] allows users to identify popular topics in blogs based on space and time. Works like [12] and [20] discover important topics in microblogs over space and time using a probabilistic approach. None of these works ,however, provides a generic algebra to undertake spatio-temporal analytics across different applications.

Google Insights (and similar Yahoo search engine log, based works[2]) allow users to identify important trends in search patterns. However, while the search keywords capture the *information seeking* behavior of the users, social media captures the *information providing* behavior of the human-sensors. The social media content captures partial user accounts, multiple perspectives and emotional state of the human users which can not be captured by a search log.

Social network based query languages (e.g. [18]), provide rich operators for analyzing the community-based (e.g. friends, groups, friend-of-a-friend) aspects of the social media data. However these tools are weak on the spatial and temporal analysis operations, which are our current focus. Spatial OLAP incorporates GIS into analytical processing of data warehouses[17], and works like [14, 15], focus on storage of moving objects and approximate answering of aggregation queries. Spatio-temporal querying over sensor networks also try to answer such questions in an energy efficient manner [5]. However, none of these works employs an image-like data representation or exploits the related benefits. Liverman et al. [10], allude to the concept of creating 'pixels' out of social science (land use pattern) data, however they come from a very different background and did not explore the use of image and video processing techniques once the representation is made. Thus, they (and the other works above) do not support spatial operations like convolution which come naturally for such a representation, nor do they allow an intuitive match between the user mental model and query operation model.

Hence, we notice significant research interest in related areas. However, to the best of our knowledge however, ours is the first attempt at creating a unified image-processing inspired architecture for undertaking spatio-temporal analysis. Our image-based representation provides a common framework for spatio-temporal data aggregation, media processing, data visualization, and an intuitive correspondence between query model and user's mental model. The generic query operators defined on top of this model allow our framework to be applicable across multiple spatio-temporal application domains.

Current version of this paper builds upon earlier versions which provided a proof-of-concept technical demonstration and an interactive poster presentation [1]. The current version provides all the relevant details, reports experimentation across multiple applications, and describes the progress in our thinking, especially underscored by a set of generic query operators which were missing in previous versions.

[1]Citation withheld for double-blind review

4. PROPOSED APPROACH

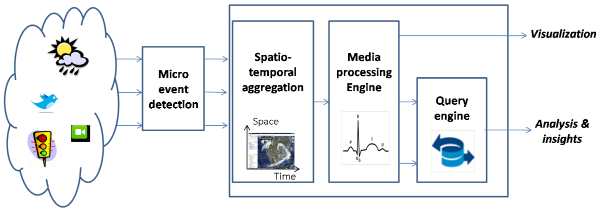

Figure 2: Overall architecture

4.1 System architecture

The overall system architecture is shown in figure 2. The system processes raw data coming from different media sources to obtain micro-level event data. It aggregates the data about any particular theme coming from a particular location into a spatio-temporal pixel (i.e. *stel*). We combine multiple such spatio-temporal blocks into *e-mage* or *temporal e-mage set* representations. Next, we undertake multiple media processing inspired operations on this data to detect important parameters, patterns and macro events as may be relevant to the application domain. These properties can then directly be visualized by the system designer to gain a global level awareness of the situation evolving, or can be used as a live data-source for users to formulate their spatio-temporal queries, a combination of which can be used for intelligent decision making.

Micro-event detection

Multiple (individual) users generate social media without explicitly coordinating with others. Our system processes such raw data to detect micro-level events which would be useful for upper level analysis. The system employs a simple bag-of-words (similar to [7]) approach to detect various micro-events from the generated media. Hence, a micro-event is detected each time a tweet containing the relevant term is posted. Similarly a micro-event is detected each time a Flickr image with matching color patterns, concepts or tags is posted.

Spatio-temporal aggregation using social pixels

The system aggregates micro-events originating from a spatio-temporal bin into a representative social pixel value. Clearly, higher level abstractions have a trade-off with lower level details. Hence, while the social pixel value assigned depends on the application, the underlying idea is to abstract away the details and maintain a representative value which would be useful for higher level processing. In the social interest modeling applications, we consider the count of micro-events related to the theme, to be this representative value. For environmental monitoring using Flickr, we use the greenery level of the images from that area to be this representation. Other plausible values include number of images with matching 'tags' or 'concepts', mean energy of audio-posts, average monthly income, demographic count, temperature info, distance to amenities, accident count etc.

Media processing engine

The 2-D grid like representation of social pixels to create e-mages, opens doors to multiple image processing operations.

The system performs multiple image and video processing operations in the background, to support the various query operators made available to the user.

Query engine

The system currently supports 6 sets of operations viz. *Selection, Arithmetic and Logical, Aggregation, Grouping, Characterization, and Pattern matching*, for analyzing any spatio-temporal data. The *Selection* operator allows the users to specify the spatio-temporal bounding box in which to perform the other operations. The *Arithmetic and Logical* operators allow the users to combine and compare multiple e-mages using operations like add, and, multiply, and convolution. The *Aggregation* operator allows temporally related e-mages to be aggregated. The *Grouping* operator allows the composed e-mage about any topic to be split into multiple e-mages on for each 'blob'/'zone' /'segment' of the e-mage. The *characterization* operators represent different attributes (e.g. epicenter, density, shape) for each of the segments. Computing the attributes separately for each segment helps to better characterize many attributes like epicenter, shape etc., which may lose their meaning if computed across multiple 'segments'/ 'clusters' as if they were one. The pattern matching can be undertaken to see how closely does the represented phenomena match any of the patterns from a library (e.g. monotonically increasing in time, explosion in interest, radial spatial growth) or related historical data.

5. DATA MODEL

In this section we present our data model which supports various query operators for the spatio-temporal data. The defined data model builds upon image algebra literature [16, 6]. While describing it as a 'social pixel' in social media setting, we generically define the fundamental building block for all the spatio-temporal-thematic data as a stel (spatio-temporal element). The *stels* are combined to create an *e-mage*, which is conceptually similar to a *multi-spectral image*, whose each band corresponds to one theme present in the data.

5.1 Stels

We consider a *stel* (spatio-temporal element), as the fundamental building block for spatio-temporal data. Each stel has its coordinates in space and time, a theme, a value, and a pointer to raw data from which the value is derived.

$$stel = [st_{coord}, theme(s), value(s), pointer(s)] \qquad (1)$$

where:

$$st_{coord} = [lat, lon, alt, timeStamp] \qquad (2)$$

5.1.1 $\mathbf{st}_{coord}$

lat, **lon**, **alt**:
Represent the geographical latitude, longitude, and altitude of the location, corresponding to the data.
$lat \in [-90, 90]$, $lon \in [-180, 180]$, $alt \in [0, \infty]$

timeStamp:
Represents the timestamp corresponding to the data. We represent timestamps in POSIX / UNIX format.
$timestamp \in [0, \infty]$

5.1.2 *Theme*

Represents the 'theme' e.g. 'Swine Flu, i-Phone, Obama, Iran-election', to which the data corresponds. This theme can be set by the application designer or defined automatically during the micro-event detection and *stel* composition.

5.1.3 *Value Set*

A value set V is an instance of a homogeneous algebra, which is a set of values and operations that manipulates these values. Here, the set of values is $\mathbb{N}$, and the operations include addition, subtraction, multiplication, division, maximum, minimum, etc.

Each stel consists of one or more [theme, value] pairs, one for each 'band' or topic being represented.

5.1.4 *Data Pointer*

This is a vector of pointers to the actual data. While we intentionally want to abstract the raw data into the above 'value', for computational and reasoning efficiency, certain application designers may want to keep pointer(s) to the actual data to support detailed analysis when required. This obviously has trade-offs with privacy of individual user, and may not be supported in all applications.

For the rest of our discussion here, we will focus on 2-D (latitude and longitude) spatial information, theme and numerical values for each *stel* and ignore the altitude component, as well as the data pointers. Also, it is often desirable to consider spatial dimension only. So we call each $[lat, lon, theme(s), value(s)]$ a 'pixel' in our discussion here.

We explicitly define three aggregates of *stels* which will be commonly used when dealing with collections of such data onto the 2-D spatial axes (*e-mage*) and spatio-temporal axes (*Temporal E-mage Set* and *Temporal Pixel Set*).

5.2 E-mage

Let V be a value set, X a two dimensional point set (lat, lon), and Θ a set of theme. An $\Theta \times V$-valued multi-thematic e-mage on X is any element of $(\Theta \times V)^X$, which is a map from X to $\Theta \times V$. Here we use grid as the data structure representation of the e-mage. An e-mage is represented as $g = (\mathbf{x}, \{(tm, v(\mathbf{x}))\} | \mathbf{x} \in X = \mathbb{R}^2$, $tm \in \Theta$ and $v(\mathbf{x}) \in V = \mathbb{N})$. We use $|g|$ to indicate the size of an e-mage, which is $(width, length)$.

5.3 Temporal E-mage Set

A temporal e-mage set, by extension is a finite set $TES = \{(t_1, g_1), ..., (t_n, g_n)\}$, where t_i is the timestamp of e-mage g_i.

5.4 Temporal Pixel Set

A temporal pixel set is $TPS = \{(t_1, p_1), ..., (t_n, p_n)\}$, where p_i is a pixel.

6. QUERY OPERATIONS

The aim of designing this operation algebra is to retrieve relevant spatio-temporal-thematical data (e-mages, TES, or their attributes) by describing their characteristics, rather than *manipulating* them directly. Just like 'Selection', 'Projection', 'Join' in relational algebra, these operators are aimed to be the basic operations, combination of which can be used for arbitrarily sophisticated querying on spatio-temporal data.

6.1 Selection Operation σ

This operator acts as a filter to choose only the subset of data which is relevant to the user application.

6.1.1 Predicate on e-mage

Predicate on e-mage g is a boolean function on pixels $P(p), p \in g$. Predicates can be applied on spatial point, theme or value of pixel p.

A spatial predicate P_R is a point lattice $R \subseteq X$. If the spatial coordinate of the pixel p is in R, $P(p)$ is evaluated to be true; otherwise, false. A theme predicate selects the $[theme, value]$ pair of the given theme P_{theme} in p. And a value predicate P_v is a value comparison on pixel value $p.v$. If the value satisfies the comparison, $P(p)$ is true; otherwise, false.

The formal definition is

$$\sigma_P(TES) = \{(t_1, \sigma(g_1)), ..., (t_n, \sigma(g_n))\} \tag{3}$$

where

$$\sigma_P(g_i) = \{(\mathbf{x}, \{(tm, y)\})\} | y = v(x) \; if \; P(\mathbf{x}, (tm, v)) \; is \; true; 0, if false\} \tag{4}$$

6.1.2 Predicate on Time

Predicate can also be applied on time of TES, which is a comparison on time point or interval. The definition is

$$\sigma_P(TES) = \{(t_{1'}, g_{1'}), ..., (t_{m'}, g_{m'})\} \; where \; P(t_{i'}) \; is \; true. \tag{5}$$

Example:

1) Show the last hour's data from TES for theme 'iPhone', from California, whose value is greater than 50.

$\sigma_{Theme=iphone \wedge R=Cal \wedge \Delta t \leq 1hr \wedge v>50}(TES)$

Note that given a multi-spectral e-mage, we can always produce a set of single-spectral e-mages by selecting the individual themes. For the ease of exposition, in the following dicussion, we assume that every e-mage in TES and every pixel in TPS is single-spectral.

6.2 Arithmetic and Logical Operation $\oplus$

Arithmetic and logical operations take two e-mage g_1 and g_2 as input, and generate a new e-mage g_3 as output. We assume that sizes of the e-mages are the same, i.e. $|g_1| = ... = |g_n|$.

$$\oplus(g_1, g_2) = g3(\mathbf{x}, \{(tm, \oplus(v_1(\mathbf{x}), v_2(\mathbf{x})))\}) \tag{6}$$

where v_1 and v_2 are the values in g_1 and g_2, and $\oplus \in \{+, -, *, /, max, min, avg, convolution, and, or\}$.

Definition of these operations can be extended to handle multiple TES, where e-mages at the corresponding timestamps in different TES are processed as follows.

$$\oplus(TES_1, ..., TES_n) = (t_1, \oplus(g_{11}, ..., g_{n1})), ..., (t_m, \oplus(g_{1m}, ..., g_{nm})) \tag{7}$$

where g_{ij} is the e-mage at timestamp j in TES_i. While most of the operations $\{+, *, max, min, avg\}$ can handle multiple inputs, some $(-, convolution)$ can handle only two inputs at a time.

6.3 Aggregation Operation α

Aggregation operation $\alpha_\oplus$ aggregates e-mages in TES based on function $\oplus$, where $\oplus \in \{+, *, avg, max, min\}$, and generates one result e-mage. Here, we assume that all e-mages $g_1, ..., g_n$ have the same size.

$$\alpha_\oplus(TES) = \oplus(g_1, ..., g_n) \tag{8}$$

Example:

1) Show the average e-mage from last one hour's e-mages of California in TES.

$\alpha_{avg}(\sigma_{R=Cal \wedge \Delta t \leq 1hr}(TES))$

6.4 Grouping Operation γ

The grouping operation groups pixels of a given e-mage g in TES based on certain function f, and slices g into a set of e-mages accordingly. The function f is a general term that expresses the function used to group the pixels.

$$\gamma_f(TES) = \gamma_f(t_1, g_1) \cup ... \cup \gamma_f(t_n, g_n) \tag{9}$$

where

$$\gamma_f(t_i, g_i) = \{g_{i1}, ..., g_{in}\} \tag{10}$$

where each g_{ij} is a sub-e-mage of g_i after applying f. For different function f, the set of result e-mages is different. We currently support $f \in \{segmentation, clustering, blob-detection\}$, for the current implementation, but generically any method can be applied to do this.

Example:

1) Segment every e-mage in the last one hour's TES for California by using segmentation.

$\gamma_{segmentation,n=3}(\sigma_{R=Cal \wedge \Delta t \leq 1hr}(TES))$

6.5 Characterization Operations: Spatial ϕ

E-mage characterization operation takes every e-mage g in TES, and computes a representative pixel to characterize this e-mage based on a function f.

$$\phi_f(TES) = \{(t_1, \phi_f(g_1)), ..., (t_n, \phi_f(g_n))\} \tag{11}$$

where $\phi_f(g_i)$ is a pixel p_i computed from $f(g_i)$. The result of this operation is a temporal pixel set TPS.

The function f can be selected from $\{count, max, min, sum, avg, epicenter, density, shape, growth-rate, periodicity\}$. For functions max, min, $epicenter$, the pixel is the one where the property of **value** is reached. For instance, after using max, the pixel whose value is the largest among all others in g is returned. However, for other functions as $count$, sum, avg, $density$, $shape$, $growth-rate$, $periodicity$, the computation produces only a value v without any specific location point. In these cases, the point of the result pixel is set as $(0, 0)$, and v is set as the pixel value.

Example:

1) Find the max point of each e-mage in last one hour's e-mages of California in TES.

$\phi_{max}(\sigma_{R=Cal \wedge \Delta t \leq 1hr}(TES))$

6.6 Characterization Operation: Temporal τ

Temporal characterization operation is designed to characterize temporal pixel set TPS. Note that TPS is the result of spatial characterization on e-mages in TES. So temporal characterizations aid users to study how the spatial characterizations *vary* over time. A prediction operation that calculates the value at next timestamp t_{k+1} based on pixels before t_{k+1} is also supported, and is treated as a special type of characterization.

$$\tau_f(TPS) = \{(t_k, f((t_1, p_1), ..., (t_k, p_k))) | k \in [2, n]\} \tag{12}$$

where $f \in \{displacement, distance, velocity, speed, acceleration, linear\ extrapolation, exponential\ growth, exponential\ decay, etc.\}$. Functions like *linear extrapolation* are used for prediction based on multiplying with appropriate kernels.

Example:

1) Find the velocity of epicenter in last one hour's e-mages of California from TES.

$\tau_{velocity}(\phi_{epicenter}(\sigma_{R=Cal \wedge \Delta t \leq 1hr}(TES)))$

S.No	Operation	Input	Output
1	Selection σ	TES	TES
2	Arithmetic $\oplus$	$K \cdot$ TES	TES
3	Aggregation α	TES	TES
4	Grouping γ	TES	TES
5	Characterization:		
	- Spatial ϕ	TES	TPS
	- Temporal τ	TPS	TPS
6	Pattern Matching:		
	- Spatial ψ	TES	TPS
	- Temporal	TPS	TPS

Table 1: Summary of various query operations. TES=Temporal E-mage Set, TPS=Temporal Pixel Set

6.7 Pattern Matching Operation ψ

Pattern matching operations compare the similarity between a TES/TPS and a pattern, which can be defined from historical data or chosen from a library of relevant patterns.

6.7.1 *Spatial Pattern Matching*

Spatial pattern matching compares every e-mage g_i in TES with a pattern e-mage P, and defines a temporal pixel set where each pixel value represents the observed similarity.

$$\psi_P(TES) = \{(t_1, p_1), ..., (t_n, p_n)\} \tag{13}$$

6.7.2 *Temporal Pattern Matching*

Similarly, we can compare the values of each pixel in TPS to certain temporal pattern P. The patterns can be monotonically increasing, decreasing, sine, cosine, etc.

$$\psi_P(TPS) = (t_n, p_n) \tag{14}$$

where value of p_n is the similarity value.

Example:
1) Compare the similarity between last one hour's e-mages of California from TES with radial decay.
$\psi_{scaled}(\sigma_{R=Cal \wedge \Delta t \leq 1hr}(TES), K_{radial\_decay})$

A summary of all the operators defined and the corresponding input and output is shown in table 1.

7. MEDIA PROCESSING FOR SUPPORTING QUERY OPERATORS

We have implemented the query operators using an underlying media processing engine. Each of the query operator corresponds to one or more classes of media processing operations. This mapping is shown in figure 3.

As can be seen, multiple query operations (e.g. 'Characterization', and 'Pattern matching') may employ the same media processing operation (e.g. 'Convolution') in terms of the underlying implementation.For example both 'Circularity' (which is a 'characteristic' from a user perspective), and Pattern matching with a library of Kernels, use convolution operation. However they are different operators from a declarative user perspective.

8. IMPLEMENTATION AND RESULTS

8.1 System implementation

S.No	Query Language Operator	Media processing Operator Category	Media processing Operator Details
1.	Selection		
	-Spatial	Arithmetic	AND with the spatial mask
	-Temporal	Arithmetic	AND with the temporal mark
	-Thematic	Arithmetic	=
	-Value	Arithmetic	AND, >, <, =
2.	Arithmetic and Logical		
	-Max, Min, +,-,%,*	Arithmetic	Max, Min, +,-,%,*
	- NOT, OR, AND,	Logical	NOT, OR, AND
	-**Convolution**	Convolution	Convolution
3.	Aggregation	Arithmetic/Logical	Max, Min, +,-,%,*, NOT, OR, AND
3.	Grouping		
	- Predefined segments count	Segmentation	K-means
	- Segment count not predefined	Segmentation	Affinity propagation
4.	Characterization		
	::Spatial		
	- Count, Min, Max, Sum, Average, Variation	Statistical	Count, Min, Max, Sum, Average
	- Coverage	Arithmetic	Count
	- Epicenter	Arithmetic	Weighted average
	- Circularity	Convolution	Scale free convolution with known circular kernel
	- Growth rate	Arithmetic	+, -, %
	:: Temporal		
	- Displacement, Distance, Velocity, Acceleration, Growth rate	Arithmetic	+, -, %, *
	- Future estimation	Arithmetic	Multiplication with Kernels based on users choice e.g. linear, progression exponential growth
	- Periodicity	Convolution	Auto correlation i.e. Self convolution with time-lagged variant.
5.	Pattern Matching		
	- Scaled Matching	Convolution	Convolution, Auto-correlation using user defined or pre-defined Kernels
	- Scale free Matching	Convolution, Statistical	Maxima from loops of Convolution/ Auto-correlation with different sizes

Figure 3: Implementation details: Mapping of each Query operator to its media processing operation

The presented analytic system has been implemented in Java. The queries defined in the query algebra are currently called as functions in Java program. For example the query 'When and where was the interest peak in our product?' corresponding to
$\tau_{max}(\phi_{max}(\sigma_{topic=P}(TES)))$ is called as
$opCharac(temporal, max, opCharac(spatial, max, opSelection(topic, P)));$

The data processed can be exported into KML files for rendering in Google Earth. A web based system-interface has also been created and a (partial) demo and datasets are available at [2]. The social media data employed is obtained from Twitter and Flickr using their respective APIs.

Twitter data was obtained using 2 sources. We use the twitter streaming API to download a portion of all public Twitter feeds. While some Twitter posts are directly GPS geo-coded, we geocoded the rest by using the 'home' location (e.g. San Fransisco, CA) of the user by using an opensource geocoding service ($http : //ws.geonames.org$). Only the tweets successfully geocoded were used for our experiments here. We augmented this data set using location based queries for each location across US for selected topics. The results presented here are based on a data corpus of more than 100 million tweets using 'Spritzer' stream (since Jun 2009), and the higher rate 'Gardenhose' stream since Nov, 2009. The data set currently gets augmented by about 1 million tweets from the live stream each day.

[2] $http : //socialemage.appspot.com/acmmm/$

S.No	Category	Event	Physical Date	Observed Temporal Peak	Physical Location	Observed Spatial Peak
1	Politics	Health Care Bill passed	2010-03-21	2010-03-21	38.89, -77.03 (Washington)	41, -74
2	Politics	California Prop 8, Trial Day 1	2010-01-11	2010-01-11	37.77, -122.41 (San Francisco)	38,-122
3	Society	Fort Hood Shootings	2009-11-05	2009-11-05	31.13, -97.78 (Fort Hood, TX)	33,-97
4	Society	SeaWorld Whale Accident	2010-02-12	2010-02-12	28.54, -81.38 (Orlando, FL)	29,-81
5	Sports	Winter Olympics Opening ceremony	2010-02-12	2010-02-12	49.24, -123.11 (Vancouver)	***44,-79***
6	Sports	Baseball World Series final	2009-11-04	2009-11-04	40.71, -74.00 (New York)	41, -74
7	Entertainment	Oscars	2010-03-07	2010-03-07	34.05, -118.24 (Los Angeles)	34, -118
8	Entertainment	South by Southwest festival	2010-03-12 to 2010-03-21	2010-03-15	30.26, -97.74 (Austin, TX)	30, -98
9	Tech. Conv.	CES 2010	2010-01-05 to 2010-01-07	2010-01-06	36.17, -115.13 (Las Vegas)	34,-118
10	Tech. Conv.	TED 2010	2010-02-10 to 2010-02-13	2010-01-10	33.76, -118.19 (Long Beach,CA)	34, -118

Figure 4: Correlation between Real world events and twitter data

The combined collection allows users to pose queries both on the current, as well as archived data. While our approach is generic, here we illustrate the results using US, as the selected spatial region.

8.2 Correlation with real world events

To verify whether social media sources do indeed capture aspects of real world events we ran the following experiment. We selected a list of 10 important events which happened in USA (or Canada) during the period of Nov 2009-Mar 2010. We tried to be diverse in both the category of events as well as the physical location. We ran the peak (in time) and max (in spatial location) operators on the Twitter corpus for these themes. To counter the effect of high Twitter user base in different parts of the country we performed 'background subtraction' (with an e-mage composed by averaging the number of posts, on the first day of each month, for *all* topics, for each location).

As shown in figure 4, we found that the temporal peak coincided with the actual event occurrence period for all 10 events. The spatial peak matched accurately for 7, and was a nearby big city for 2 more, out of 10 events. Only for the Winter Olympics event, the peak (Toronto) was not within reasonable distance of the original location (Vancouver).

While a more detailed analysis is required to understand such phenomena, we found the results to highlight reasonable correlation between real world events and their manifestation on social media data.

8.3 Application: Business analysis

While the spatio-temporal event detection experiment in sec 8.2 only dealt with single events (with single spatial and temporal peaks), we need more sophisticated tools to analyze spatio-temporal situations which may have multiple peaks, locations, and may travel across space and time. Here, we demonstrate visualization and sample queries which might be asked by a business analyst, when dealing with spatio-temporal data about their product of interest (P) in TES (e.g. 'iPhone').

8.3.1 Visualization for situation awareness

To support easy situation awareness we visualized the data about newly launched iPhone [3], in Google Earth (see figure 5). We analyzed the tweets and added representative icons to represent information about areas from where the users felt that the 'iPhone' was 'too expensive', 'reasonably priced', showed an interest in 'other service providers', or 'Unlocking of iPhone', and were 'happy' or 'unhappy' with the services of $AT\&T$. Each of these layers of information could be switched on and off. Note that while the added icons do not add any new information from an information theoretic perspective, they can significantly improve the situation awareness for any human user.

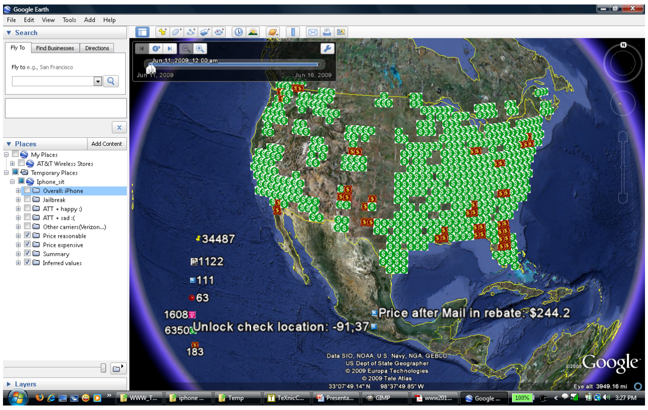

Figure 5: Visualization in Google Earth of different aspects of the tweets about 'iPhone'

8.3.2 Query engine for analysis and insights

Next, we consider some sample queries.

1. When did the interest peak in our product?
$\tau_{max}(\phi_{sum}(\sigma_{R=USA \wedge theme=P}(TES)))$
2. Show me three different zones for interest patterns about our product.
$\gamma_{segmentation,n=3}(\sigma_{R=USA \wedge theme=P}(TES))$
3. Where are the epicenters for each 'zone' of product interest?
$\phi_{epicenter}(\gamma_{n=3}(\sigma_{R=USA \wedge theme=P}(TES)))$
4. What is the overall number of users *happy* with our product?
$\phi_{sum}(\alpha_{+}(\oplus_{AND}(\sigma_{R=USA \wedge theme=P \wedge \Delta T=7}(TES)),$
$\sigma_{R=USA \wedge theme=\text{'happy'} \wedge \Delta T=7}(TES))))$
5. Assuming linear extrapolation, show me the total anticipated interest tomorrow (Jun16 for this example)?
$\tau_{linear-extrapolation}(\phi_{sum}(\sigma_{R=USA \wedge theme=P \wedge \Delta T=7}(TES)))$
6. Show me the best location to open a new store for product P, given existing store locations e-mage S, and 'catchment' area kernel for each store C.
$\phi_{max}(\oplus_{Convolution}(\oplus_{-}(\alpha_{+}(\sigma_{R=USA \wedge theme=P}(TES)),$
$\oplus_{Convolution}\ (\sigma_{R=USA}(S), C)), C))$

The data considered for answering these queries was from Jun 2 to Jun 15, 2009, which included the date of iPhone 3G version's release (Jun 8). The results of applying queries 1 through 3 have been shown in Fig. 6. As can be noticed (amongst other things), the peak of interest does indeed match with the actual release date. The different zones

[3] data from Jun 2009

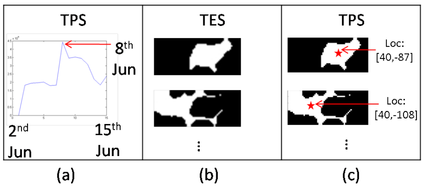

Figure 6: Results of applying the queries 1-3 on Twitter data for 'iPhone' theme

(query 2), and epicenters found (query 3) also make reasonable sense. The overall number of 'happy' users[4] was found to be 1297 (query 4), and anticipated interest (query 5) was found to be 29120.

For easier understanding we visually illustrate (see fig. 7) query 6 above and show a sample result. The query aims to find the best location to start a new store for 'iPhone'. In this example, e-mages corresponding to the term 'iPhone' were *aggregated* for 14 days. On the other hand, the *AT&T* retail store[5] locations e-mage was *convolved* with the catchment area (assumed to be a Gaussian Kernel C) for each store. The difference between the obtained 'aggregate interest' and 'net catchment area' images was taken to be a representative of the 'under-served interest areas' where it makes sense to open a new retail store. To find out the best location for such a store, we undertook a *convolution* operation between the catchment area of a new store and this 'under-served interest areas' image. The *maxima* operation on the obtained image gave the most appropriate location. Note that the obtained answer in this case is not merely a pixel location but a physical geo-location with real life semantics. We obtained the details of the physical location via reverse-geocoding.

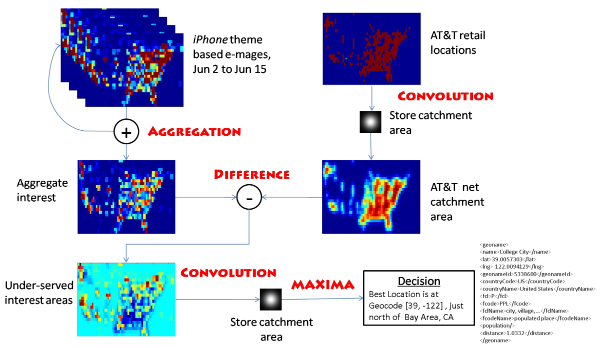

Figure 7: Combination of operators for undertaking business decision

8.4 Application: Political event analytics

8.4.1 Query engine for analysis and insights

We demonstrate sample queries which might be asked by a political analyst, or campaign manager, when dealing with spatio-temporal data about a personality of interest (P) (e.g. 'Obama') or issue of interest (I) (e.g. 'Healthcare').

[4]detected by simply looking up for ':)' in text for now.

[5]*AT&T* is the only legal service provider for iPhones in US. For the purpose of our discussion, we will assume that iPhones can only be sold from these retail locations.

1. When did the interest peak about personality P?
   $\tau_{max}(\phi_{sum}(\sigma_{R=USA\wedge theme=P\wedge \Delta T=120days}(TES)))$
2. What is the periodicity of the interest in this personality?
   $\tau_{periodicity}(\phi_{sum}(\sigma_{R=USA\wedge theme=P\wedge \Delta T=120days}(TES)))$
3. Show me the trajectory for interest in personality P over last 7 days?
   $\phi_{Epicenter}(\sigma_{R=USA\wedge theme=P\wedge \Delta T=7days}(TES))$
4. Show me the aggregate interest in P over last 7 days, in Republican states ?
   $\oplus_{AND}(\alpha_{+}(\sigma_{R=USA\wedge theme=P\wedge \Delta T=7days}(TES)),$
   $\sigma_{R=USA\wedge theme=Republican\wedge \Delta T=7days}(TES))$
5. Show me the interest in issue I, when interest in personality P gained its peak?
   $\sigma_{theme=I\wedge R=USA\wedge time=tp}(TES)$
   where $tp = \tau_{max}(\phi_{sum}(\sigma_{R=USA\wedge theme=P\wedge \Delta T=120days}(TES)))$
6. What is the similarity between the interest patterns for I and P on the above date?
   $\psi_{spatial-scaled-match}(\sigma_{theme=I\wedge R=USA\wedge time=tp}(TES),$
   $\sigma_{theme=P\wedge R=USA\wedge time=tp}(TES))$
   where $tp = \tau_{max}(\phi_{sum}(\sigma_{R=USA\wedge theme=P\wedge \Delta T=120days}(TES)))$

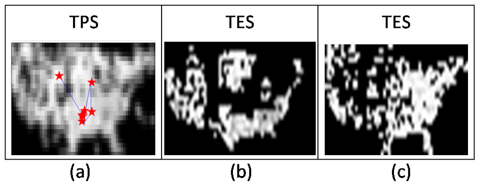

Figure 8: Results for queries 3-5 for politician P='Obama' and issue I='Healthcare'

The results presented here are based on data collected for Personality 'Obama' and issue 'Healthcare' between Nov 5th 2009, and Mar 30, 2010. The peak of interest in 'Obama' (Query 1) was found on 27th Jan, 2010, which corresponds with the 'State of the Union' address. Obviously multiple events of interests (and local maximas) about 'Obama' occurred during the 4 months. The periodicity value (query 2) was found to be approx. 20 days. The trajectory (query 3) of interest epicenter for 7 days (Jan 21-Jan 28) is shown in Fig. 8. Query 4, i.e. the interest in Republican states (based on current governorship in the state), and the e-mage for 'Healthcare' on the peak day (query 5) are also shown in Fig. 8. Lastly, the similarity value (query 6) between interest in 'Obama' and 'Healthcare' for that day was found to be 0.693.

8.5 Application: Seasonal characteristics analysis

Our experiments so far have focused on using number of twitter posts from a particular location as the social pixel or stel value. However our approach is generic. In this section we consider average green color intensity [6] of all Flickr images uploaded from a location to be the representative stel value for that location. The experiment is aimed at analyzing climatic/ seasonal phenomena as they occur across different parts of US. Specifically we want to see which parts

[6]based on normalized data using $value = G - average(R + G + B)$

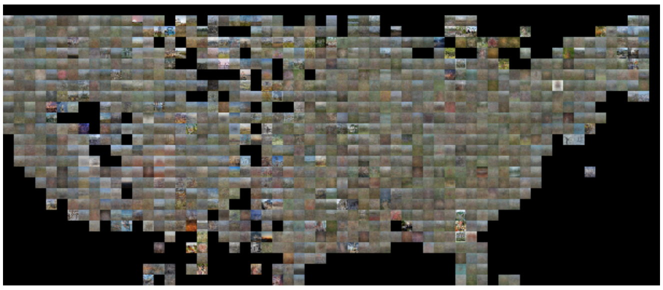

Figure 9: Spatio-(temporal) emage representing average of Flickr images posted from each location for the month of Aug 2009.

of US are more green than others, and what are the seasonal patterns or trajectory of such phenomena. We envision such an approach getting extended to detecting when certain phenomena occur (e.g. Fall colors appear in New York), or even understanding flora growth, or bird migration patterns, by using more sophisticated concept detectors for images in future. Here, we show sample queries for greenery based on color intensity, and 'snow' based on number of images with matching meta-data (i.e. tags).

1. Show me the variation in green color intensity as it varies over the whole year.
$\phi_{sum}(\sigma_{theme=green \wedge \Delta t=1yr}(TES))$

2. Where is the peak in the greenery across whole of US?
$\phi_{max}(\alpha_{+}(\sigma_{theme=green \wedge \Delta t=1yr}(TES)))$

3. Show me three segments based on greenery, as they vary over the year.
$\gamma_{segmentation,n=3}(\sigma_{theme=green \wedge \Delta t=1yr}(TES))$

4. Show me difference between red and green colors, for the New England region as it varies over the year.
$\oplus_{-}(\phi_{sum}(\sigma_{theme=red \wedge R=[(40,-76),(44,-71)] \wedge \Delta t=1yr}(TES)),$
$\phi_{sum}(\sigma_{theme=green \wedge R=[(40,-76),(44,-71)] \wedge \Delta t=1yr}(TES)))$

5. Show me the trajectory of the epicenter of high-snow activity region throughout the year.
$\phi_{epicenter}(\sigma_{val=1}(\gamma_{n=3}(\sigma_{theme=snow \wedge \Delta t=1yr}(TES))))$

6. What is the degree of similarity between aggregated snow activity and the North to South linear decay pattern?
$\psi_{scale-free}(\alpha_{+}(\sigma_{theme=snow \wedge \Delta t=1yr}(TES)), K_{north-south-decay})$

The results here are based on TES created from Flickr data from US, at 1 month granularity for the year 2009. Image processing (i.e. averaging of pixel intensities) was restricted to first 100 samples from each location. Hence, a total of 706,415 images were aggregated for answering queries 1-4. A sample social image created by averaging images from all over US is shown in figure 9. As shown in Fig. 10, the overall green color intensity (query 1) peaked during summer months. The area with most green pictures (query 2) was at [35,-84], which happens to be at the intersection of 3 national forests and 1 national park. The overall variation in zones of different greenery showed a reasonable trend, with high greenery zone moving from south-east of US in April towards north, covering most of US in summers and then receding southwards again. The relative intensity of red and green showed interesting trends. For example, in New England region (query 4), the green dominated red over summer months, but red overtook green during the 'Fall' months.

The queries 5 and 6 were answered using meta-data from 79,628 images. The 'snow' tags data was normalized using the number of geocoded images uploaded from that region on any topic. The trajectory for the epicenter of the high snow segment is shown in Fig.10(e). Lastly, the similarity value between the aggregate snow activity and north-south linear decay pattern was found to be 0.519 .

8.6 Discussion and future work

Our experiments first verified the correspondence between real world events and social media (sec 8.2). Going beyond single event detection, sections 8.3, 8.4 demonstrated the *expressiveness* of the defined query algebra to pose spatio-temporal situation queries which can be answered using sophisticated media processing tools under the hood. The experiments clearly show the value of moving beyond single event detection to handling trajectories, periodicity, future state estimation, correlation across topics, spatio-temporal visualization, and similarity estimates. Admittedly, the results for different applications are preliminary in terms of data granularity, and processing employed. But they clearly show the potential, and the hence the need for such a spatio-temporal query algebra for undertaking more detailed analysis. Application results are only going to get better with more availability of fine-grained geo-data, and more sophisticated analysis techniques, both of which are becoming increasingly better each day.

Experiment 4 (sec 8.5) demonstrated that the proposed approach can work across media sources and support different social pixel representations and applications. Together, the experiments demonstrate a working system involving a large corpus (>100 million tweets, and >700,000 Flickr images) to highlight the 3 design principles set forth in section 2. The use of human sensors, a social pixel approach, and the situational query algebra are indeed demonstrated to be useful in multiple application domains.

This work is aimed to be a foundation for multiple research challenges and opportunities ahead. We kept the following problems out of the scope of the current paper, but intend to work and collaborate on them in near future.
1) Defining a (visual) query language on top of the query algebra defined.
2) User evaluation from multiple spatio-temporal domains and operator refinement.
3) Automatic topic modeling for thematic analysis.
4) Creating reverse-911 like control/ recommendation applications based on reasoning presented.
5) While we have focused in this work on social media and human-sensors, the techniques extend seamlessly to all kinds of sensors. In fact, we expect the distinction to get increasingly blurred with time. Growth of multimodal, sensor based microblogs and works like house that tweets [7] are steps in that direction. We intend to evaluate our work on multiple sensor platforms including automatic device tweets, traffic lights, weather info in near future.

[7] $http://asmarterplanet.com/blog/2009/07/a-house-that-tweets.html$

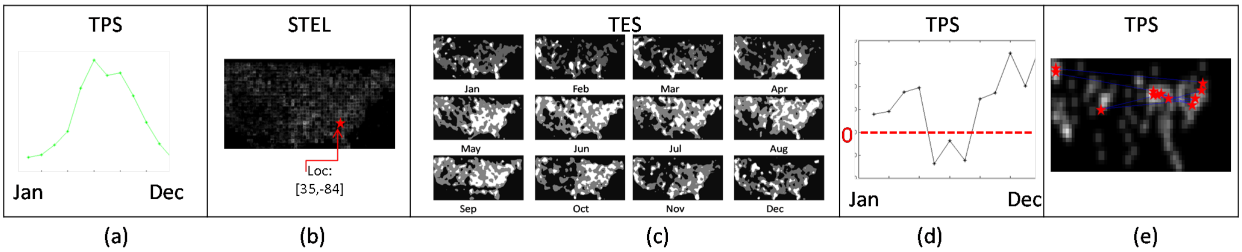

Figure 10: Answers for queries 1-5 for seasonal characteristics monitoring using Flickr data

9. CONCLUSIONS

In this work, we have presented a case for using human sensors to detect real world events, and generate situation awareness. We have described how spatio-temporal-thematic data in various social media can be aggregated into 'social pixels'. An image like representation allows for sophisticated data processing, but the implementation details cna be hidden from a user, who simply employs a declarative query algebra to pose relevant queries. The designed operators can be combined to define arbitrarily sophisticated situational queries. Results of applying this approach across multiple applications have been demonstrated using a growing corpus containing millions of user posts. We intend to extend the approach to different types (sources and modalities) of social media, and implement a full visual query language in near future.

10. REFERENCES

[1] R. Albert, H. Jeong, and A. Barabási. The diameter of the world wide web. *Nature*, 1999.

[2] L. Backstrom, J. Kleinberg, R. Kumar, and J. Novak. Spatial variation in search engine queries. In *World Wide Web conference*, 2008.

[3] N. Bansal and N. Koudas. Blogscope: a system for online analysis of high volume text streams. In *PVLDB*, pages 1410–1413, 2007.

[4] J. Biehl, M. Czerwinski, G. Smith, and G. Robertson. FASTDash: a visual dashboard for fostering awareness in software teams. In *Proc. CHI*, 2007.

[5] A. Coman, M. Nascimento, and J. Sander. A framework for spatio-temporal query processing over wireless sensor networks. In *VLDB: workshop on Data management for sensor networks*, 2004.

[6] Q. Hart and M. Gertz. Optimization of multiple continuous queries over streaming satellite data. In *ACM GIS conference*, 2006.

[7] L. Hirschman, M. Light, E. Breck, and J. Burger. Deep Read: A reading comprehension system. In *ACL Computational Linguistics*, 1999.

[8] N. Jacobs, W. Burgin, N. Fridrich, A. Abrams, K. Miskell, B. Braswell, A. Richardson, and R. Pless. The global network of outdoor webcams: properties and applications. In *ACM GIS conference*, 2009.

[9] S. Kingsbury. Wisdom for the Masses. *American Speech*, pages 358–360, 1987.

[10] D. Liverman. *People and pixels: Linking remote sensing and social science*. Natl Academy Pr, 1998.

[11] J. Lloyd. Practical advantages of declarative programming. In *Joint Conference on Declarative Programming, GULP-PRODE*, volume 94, 1994.

[12] Q. Mei, C. Liu, H. Su, and C. Zhai. A probabilistic approach to spatiotemporal theme pattern mining on weblogs. In *World Wide Web Conference*, 2006.

[13] M. Nagarajan, K. Gomadam, A. Sheth, A. Ranabahu, R. Mutharaju, and A. Jadhav. Spatio-temporal-thematic analysis of citizen sensor data: Challenges and experiences. *Proc. WISE*, 2009.

[14] S. Orlando, R. Orsini, A. Raffaetà, A. Roncato, and C. Silvestri. Spatio-temporal aggregations in trajectory data warehouses. *Data Warehousing and Knowledge Discovery*, 2007.

[15] N. Pelekis, Y. Theodoridis, S. Vosinakis, and T. Panayiotopoulos. Hermes - a framework for location-based data management. In *Proc. EDBT*, 2006.

[16] G. X. Ritter and J. N. Wilson. *Handbook of Computer Vision Algorithms in Image Algebra*. CRC Press, Inc., Boca Raton, FL, USA, 2000.

[17] S. Rivest, Y. BÃl'dard, and P. March. Towards better support for spatial decision-making: defining the characteristics. In *Geomatica*, pages 539–555, 2001.

[18] R. Ronen and O. Shmueli. SoQL: A Language for Querying and Creating Data in Social Networks. In *ICDE*, pages 1595–1602, 2009.

[19] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors. In *World Wide Web Conference*, 2010.

[20] J. Sankaranarayanan, H. Samet, B. Teitler, M. Lieberman, and J. Sperling. TwitterStand: news in tweets. In *ACM GIS conference*, 2009.

[21] D. Shamma, L. Kennedy, and E. Churchill. Tweet the debates: understanding community annotation of uncollected sources. In *SIGMM-Workshop on Social Media*, 2009.

[22] A. Sheth. Citizen sensing, social signals, and enriching human experience. *IEEE Internet Computing*, 13(4):87–92, 2009.

[23] A. Signorini. Swine Flu monitoring using twitter. In $http://compepi.cs.uiowa.edu/\ alessio/twitter-monitor-swine-flu/$.

[24] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from Internet photo collections. *International Journal of Computer Vision*, 80(2), November 2008.

[25] J. D. Ullman. *Principles of Database and Knowledge-Base Systems, Volume I*. Computer Science Press, 1988.