Title
Magic of Numbers in Networks of Wireless Image Sensors

Authors
Rahimi, Mohammed
Ahmadian, Shaun
Zats, David
et al.

Publication Date
2006-10-31

Peer reviewed

Magic of Numbers in Networks of Wireless Image Sensors

Mohammad Rahimi, Shaun Ahmadian, David Zats, Rafael Laufer, Deborah Estrin
Center for Embedded Networked Sensing
University of California at Los Angeles

mhr@cens.ucla.edu, ahmadian@lecs.cs.ucla.edu,
dzats@ucla.edu, rlaufer@cs.ucla.edu,destrin@cs.ucla.edu

Abstract

Large-scale networks of battery-operated wireless image sensors have become technologically feasible. However, it is still unclear how we can benefit from large-scale deployments of imagers. In this paper, we argue that using a large number of low-power image sensors is useful and necessary in many cases. For instance, occluded environments cannot be efficiently observed with a small number of cameras. In this case, distributed imagers can provide better coverage due to minimum infrastructure requirements and availability in large numbers. Additional benefits, such as pose diversity, statistical advantages, and multiple perspectives are discussed in detail using application examples and qualitative arguments.

Keywords

Distributed Imaging, Energy Consumption, Data Size, Complexity, Computation

1 Introduction

Distributed vision systems have been extensively studied in the field of computer vision and robotics. However, the current use of imaging has been mostly restricted to resource-rich conventional cameras, such as webcams that require capable computers and permanent power sources. We have witnessed many instances of such systems exploiting deployed communication infrastructures, such as local area networks, wide area networks, and the Internet [1].

The emergence of large-scale wireless image sensors introduces new challenges and opportunities. The distinguishing property of these networks is the emphasis on low-power battery-operated devices [3].On one hand, these systems impose several resource constraints, such as memory and processing limitations. On the other hand, they give us the ability to deploy imagers in large numbers and in many different locations. This, in turn, gives birth to the idea of wireless sensor networks with dozens or hundreds of image sensors at dense spatial settings, all capable of acquiring and locally processing the images.

The are many application potentials for these networks. They includes a range of applications that exploit analysis of color, for example, monitoring the status of crops in farmlands or wineries and monitoring plants in biology experiments. Additionally, we have security as perhaps the most promising application where cameras can be used in very dense settings to monitor any unauthorized access to the environment or displacement of valuable objects. For example, cameras can be used to monitor objects such as art pieces in galleries or merchandise in shopping malls.

Due to limited available bandwidth in a large scale multihop wireless network, nodes should process the images locally in order to reduce the per-node bandwidth usage in the network. The locally extracted meta-data from the images must then be collected and aggregated to present the results to the imaging application. However, conventional image analysis and computing utilizes heavy computation for estimating properties of the images which is currently beyond the capabilities of embedded imaging devices.

In this paper, we argue that by leveraging the large number of observations in large-scale imaging network, we improve the performance of the image sensing applications in occluded environments. In these environments, several obstacles prevent the camera from seeing the intended objects. As a result, a high-resolution image is ineffective at the processing stage leading to additional energy consumption. Therefore, we argue for a reduction of both the input image resolution and processing complexity on each individual image sensor node. This, in turn, relaxes the power requirements on each node and extends the longevity of the network. The rest of the paper will elaborate these topics in the context of concrete examples.

This rest of the paper is organized as follows. In Section 2, we analyze how a larger image processing problem can be divided in smaller problems and how the energy of the overall processing is affected. In Section 3, we define when it is possible or justifiable to perform such a decomposition. Section 4 explains how the large number of sensors can be useful in many different cases, such as occluded and sparse environments. Finally, Section 5 concludes the paper.

2 Problem Statement

In deploying a network of image sensors, we essentially decompose an image sensing problem into a set of smaller subproblems, as seen in Figure 1. Like a divide and conquer approach, each subproblem is solved individually and the results are combined for the final solution. Take for example the counting problem, where we are interested in counting the number of people in a public place. We can divide this problem into local counting subproblems that are solved individually and further aggregated to solve the global counting problem. Other examples include measurement of flow

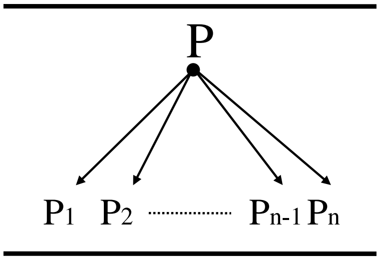

Figure 1. In deploying a network of image sensors, we essentially divide a global image sensing problem into a set of smaller image sensing subproblems. In our discussion, we assume that the boundary of subproblems are individual image sensor nodes.

of people in a public space or displacement of important assets in a building. These problems are global problems that can be solved locally. For instance, by measuring the local flow of people we can combine the results to produce a map of the flow of the population in a building, or by monitoring assets locally we can find displacement of any valuable object in the building. In general, we can write the amount of energy consumed to solve the global problem of interest as:

$$\mathrm{E}_{problem} = E_{aggregation} + \sum_{i=1}^{n} E_{subproblem} \qquad (1)$$

which includes the cost of solving the collection of subproblems and aggregating the results at some destination node. Throughout this paper, we assume that the boundaries of subproblems in such a system are the individual image sensor nodes (in Section 3 we show why that is the case). In addition, we assume that subproblems are solved on the battery-operated image sensor nodes and the aggregation of local results happens on a resource-rich destination node with access to permanent energy supply. Hence, we focus our attention on the amount of energy that each individual image sensor node consumes to locally solve a subproblem. The energy cost of solving each subproblem is:

$$\mathrm{E}_{subproblem} = E_{aquisition} + E_{communication} + E_{computation} \qquad (2)$$

which is the energy cost of acquiring the local image(s) , computing the local result on the image sensor and transmitting the result to a destination node. In this equation, the energy cost of acquisition of an image often grows linearly with the size of the input image and mainly depends on acquisition technology. However, the energy cost of communication is a complex function that depends on the size and frequency of the outgoing computed meta-data and the number of incoming and outgoing messages to route traffic in a multi-hop network [2].

Although the communication aspect is an interesting problem in itself, in this paper we want to address a less-studied component in sensor networks: local computation for image processing. Up to this point much of the research on in-network aggregation has dealt with low data rate sensors such as temperature and humidity [7]. The limited research that has been done on high data rate sensors like acoustic sensors have been fruitful in showing the need for local computation.

In Equation 2, the last term is the energy cost of running computation on the incoming images locally. This depends on the image size and often grows faster than linear, which emphasizes the importance of image computation in determining the lifetime of the image sensor. In Table 1, we see some typical image processing algorithms and their respective complexities.

Table 1. Typical algorithms' complexities [6].

Algorithm	Complexity
Classical 1-D DCT	$O(n \log n)$
Classical 2-D DCT	$O(n^2 \log n)$
Wavelet packet compression	$O(n^2 \log n)$
FFT	$O(n \log n)$

For instance, let's assume that the algorithmic complexity of an image processing algorithm is $O(r^2 \log r)$, where r is the resolution of the image in pixels. In this case, using a lower resolution considerably reduces the computation cost. We argue that by leveraging the availability of large number of sensors and minimal infrastructure constraint, we can reduce the complexity and image resolution requirements in solving each individual subproblem. The rest of this paper addresses this in more depth.

3 Justification for Problem Decomposition

In general, two factors influence the resolution of the image. First, the environment in which cameras are deployed affects the choice of image resolution. For instance, in a very occluded environment, a high-resolution camera wastes energy without much benefit. As a result, the amount of obstacles in the environment imposes a maximum bound on the image resolution.The second factor that affects the image resolution is the application itself. Several applications, such as an automatic lip reader, usually require a minimum resolution to perform properly.

In the ensuing sections, we argue as to why these two factors can result in reducing the requirement of the image resolution of each sensor while deploying a larger set of image sensors. Given that it is best to distribute resources, there is still the question of where and how we process the images. The simplest solution would be to send back all the images captured by the distributed low-resolution cameras and process them centrally. Unfortunately as Gupta *et al.* [2] demonstrated, the bandwidth of any node in a typical multi-hop network is reduced by a factor of $\sqrt{n \log(n)}$ with n being the total number of nodes. This is in fact a theoretical upper bound and in practical cases the bandwidth limitation is even worse. The bandwidth constraints force us to move the computation from the central server to the individual sensor nodes. Furthermore, these computations should be confined to single devices and not span over a group because of network bandwidth limitations, communication overhead, and resource constraints on the devices.

4 Numbers Come to Our Rescue

In this section, we present different reasons to believe that a large number of low-resolution image sensors can be ben-

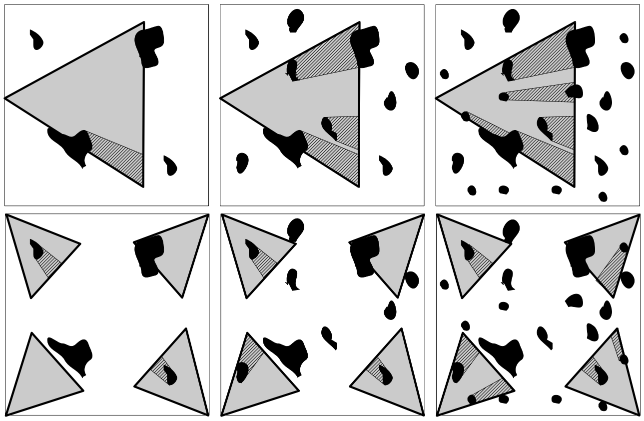

Figure 2. **Top row is a two dimensional representation of environment with a high resolution sensor. Bottom row is a group of four reduced resolution sensors with the same spatial coverage. Gray areas are the visible regions by the sensors, solid areas are infra-structure objects in the environment, and hatched areas are occluded regions.**

eficial in special cases. First, one can take advantage of multiple imagers in occluded environments where a single high-resolution camera cannot get a clear view of the regions of interest. Second, since we do not depend on the available power and communication infrastructure with wireless battery-operated image sensors, we have greater flexibility to adjust their poses . Third, in some cases, a large set of observations contributes to more accurate statistics about the environment. Finally, multiple perspectives of the same object can help in different applications, such as object classification. We now explore each of the aforementioned cases individually.

4.1 Sensing Visibility

In a practical deployment scenario, we can often determine the resolution of the image sensors based on the degree of visibility in the environment. In an unobstructed environment with high degree of visibility, we deploy high resolution sensors to cover regions of interest. However, as the volume of the objects in the environment increases, the coverage performance of a higher resolution sensor degrades due to the effect of the occlusion in the environment.

Figure 2 illustrates the effect of occlusions on the coverage of the image sensors. In this figure, the top row is a two-dimensional representation of the environment which is covered by a high-resolution sensor and the bottom row is the same environment covered with four reduced-resolution sensors. In both cases, the total area covered by the sensors are equal in the absence of occlusion. Figure 2 illustrates that as the number of occluding objects in the environment increases, the coverage of the higher-resolution sensor decreases more severely. This is because objects occlude large regions of visibility in a high-resolution image sensor.

In general, finding a global optimum value for the resolution of the image sensors based on the amount of occlusion in the environment is a difficult problem and many variant of this being studied in the field of computational geometry [4]. However, we observe that a collection of reduced quality sensors outperforms a single sensor with equivalent coverage as the degree of occlusion in the environment increases. We attribute the adequacy of low resolution cameras to the fact that as the number of occlusions increases, the maximum sensing distance decreases. As we will see later in the section, resolution translates into depth of sensing. Thus, if we are limited by occlusions, we can reduce resolution and therefore power consumption while maintaining sensing quality.

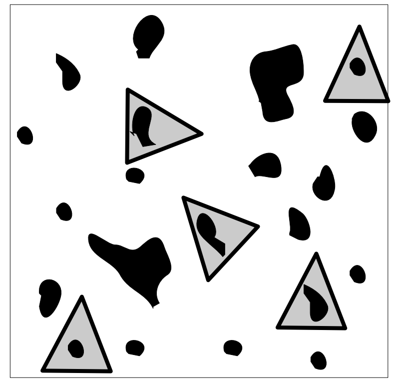

Figure 3. This figure illustrates the effect of close observation of the objects on relaxing the requirement of each sensor's resolution.

4.2 Pose Diversity

We can benefit from multiple imagers by placing them in many different poses. A pose is the combination of both the position of the camera and its orientation (i.e., where it is pointed at). Due to limited requirements for infrastructure, we can adjust the poses of the image sensors to 1) minimize the effects of the environmental variations such as illumination condition and shadows and 2) improve the quality of sensing in the regions of interest. This otherwise, in limited number of high quality images, requires heavy processing of the images to compensate for such variations [5]. We can divide the benefits of the pose diversity in two components of flexibility in relative distances and relative orientation as described below.

4.2.1 Relative Distance

In many applications, the goal of the experiment is partial coverage of the environment such as monitoring access to a building (e.g., doors, windows) or important objects of interest. In these cases, sensors can be mounted in close proximity to the regions of interest to reduce each individual sensor coverage requirement. Figure 3 depicts a network of battery-operated image sensors placed in close proximity to the objects of interest. This special arrangement in turn has several beneficial consequences in terms of power consumption as well as complexity.

Geometrically speaking, the resolution of an image sensor determines the extent of its spatial coverage. A higher resolution sensor covers objects in the space with greater spatial detail, resulting in increased spatial coverage. Let us assume that the maximum acceptable representation of a pixel is d^2 units of area in the "real world," as seen in Figure 4. Then,

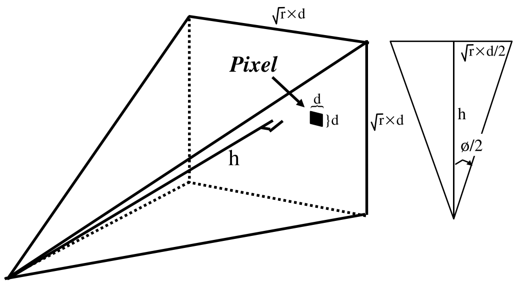

Figure 4. This figure illustrates a pyramid model of an image sensor where the minimum of a pixel's spatial representation is d unit of distance. Right is top view of the pyramid.

the space that the image sensor covers can be modeled as the volume of a pyramid. If the resolution of the sensor is r pixels, then each side of the base of such pyramid[1] is $\sqrt{r} \times d$ hence the area of its base is $r \times d^2$. The height of the pyramid can be determined based on the *Field Of View* (ϕ) of the sensor. We can write the space that is covered by the sensor as:

$$\mathrm{V}_{coverage} = \frac{1}{3} \times \frac{\sqrt{r} \times d}{2\tan(\phi/2)} \times (r \times d^2) \tag{3}$$

If we factor constants in Equation 3, the coverage model of the sensor can be written as:

$$\mathrm{V}_{coverage} = K_c \times r^{3/2} \tag{4}$$

where K_c is a constant factor. The growth of the coverage with respect to the image sensor resolution is faster than linear in Equation 4, indicating a clear benefit in enhancing the image resolution. On the other hand, a larger resolution produces a larger data size and demands additional computation time. While the computation time is less critical in devices with permanent power access, it is crucial in battery-operated devices. If we assume that the amount of time that it takes for the sensing computation to run on each image sensor is $O(\rho)$, we can write the total amount of energy consumed on an image sensor to be:

$$\mathrm{E}_{computation} = P_{computation} \times O(\rho) \tag{5}$$

where $P_{computation}$ is the amount of power that the sensor consumes during the computation state. Equation 5 suggests that, for a class of image processing algorithms whose complexity is $\rho > r^{3/2}$, the total computation energy grows faster than the coverage of the environment in Equation 4.This further highlights the importance of reduction of the image resolution to the extent that the coverage requirement in the application warrants it.

4.2.2 Relative Orientation

In many computer vision problems, the pose of the image sensors relative to the objects of interest in the environment plays a significant role in determining the complexity of the problem. This is due to view-dependent properties of the objects. Since small battery-operated image sensors require minimum infrastructure, they can be mounted in many locations and in variety of different poses. For example, we can set the orientation of the image sensors such that they view the objects of interest from a more distinctive perspective. In addition, the orientation setting can also be used to minimize the possibility of objects being partially or completely occluded.

[1] Assuming aspect ratio of the sensor to be one.

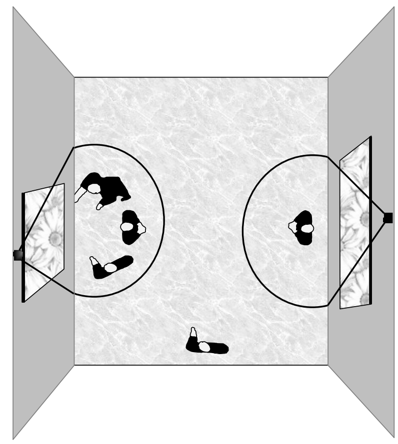

Figure 5. An image sensor mounted on top of art pieces to measure number of local visitors.

Figure 5 illustrates this concept. Consider a problem where we want to periodically estimate the instantaneous number of people and their spatial distribution in an environment such as an art museum or a public building. In this case, we can choose to mount the image sensors on the ceiling, looking from the top towards the floor. The advantage of ceiling mounted sensors is minimizing the possibility of humans occluding each other. Additionally, looking from the top, humans have highly distinctive geometrical properties and as a result we can use simpler detection algorithms.

4.3 Statistical Convergence

There are many applications where we use a network of image sensors to determine the gross statistics of the environment. In many of these cases, a large number of low-power battery-operated sensors contribute to the availability of a large spatially-distributed observation set. This plays a significant role in relaxing the requirements of sensing precision in each node. In addition, it contributes to the reduction of complexity of the subproblems on the node, to the extent that it still satisfies the precision requirement of the global problem.

Take for instance, the problem of estimating the instantaneous number of people in a public place. In this case, we perform local observations by deploying sensors in selected regions of interest to estimate the number of nearby people and predict the true value of the number of people in the entire space. We can write the aggregation function as:

$$C = \frac{A}{a} \times \frac{\sum_{i=1}^{N}(c_i + \varepsilon_i)}{N} \tag{6}$$

where C is the estimate of the total number of people in the environment. The parameter A is the total area of the environment, a is the area covered by each individual sensor and

N is the number of sensors in the environment. c_i is the result of the local counting subproblem, which is synchronously performed by the image sensors and ε_i is the amount of error in solving each subproblem. Here, we assume that sensors do not have a common field of view. In addition, we assume that due to limited visibility in the space, such as indoor buildings or crowded public spaces where people occlude each other, we do not benefit from using high resolution image sensors.

The error in Equation 6 consists of error components. The first is the error in the estimation due to limited sampling of the environment (not shown in Equation 6) and the second is the accumulative measurement error due to inaccuracies in the underlying sensing computation (i.e. $\frac{A}{a} \times \frac{\sum_{i=1}^{N}(\varepsilon_i)}{N}$). If we assume that the spatial observations of sensors are independent and identically distributed, then the former component of error is inversely proportional to the number of observations N. In addition, if we assume that local image computation error (i.e, ε_i) is a zero-mean random variable, the amount of this error will converge towards zero as the number of observations increases in a large-scale network.

4.4 Multiple Perspective

In a broad class of problems, we can take advantage of view-dependent properties of the objects by deploying sensors that look at the objects from different perspectives, as shown in Figure 6. In this case, we essentially break subproblems into smaller problem units by exploiting a number of image sensors performing computation in concert. Take for instance an object classification problem, where we are interested in classifying human intruders in the perimeter of a building. In this case, we can exploit pairs of sensors to classify humans regarding other intruding objects, such as birds, animals, and volatile projection of shadows in the outdoor environment. The clear advantage of a pair of sensors (e.g., one looking from the top and one from the side) is the ability to classify humans with much higher degree of confidence by assessment of their geometrical properties.

From the energy perspective, using multiple image sensors is justifiable if we can take advantage of multiple views to decompose a subproblem into smaller units such that

$$E_{p_{11}} + E_{p_{12}} << E_{p_1}, \qquad (7)$$

where p_1 is the initial subproblem and $E_{p_{11}}$ and $E_{p_{12}}$ are the smaller subproblem units. For this to work, the total energy cost of smaller subproblem units must be sufficiently smaller than the initial subproblem to cover the overhead of communication cost. In practice however, by sequentially executing the smaller subproblem units we can save energy by avoiding unnecessary computation at the same time that we enhance the quality of sensing.

5 Concluding Remarks

In a network of battery-operated image sensors, limited availability of power resources poses a constraint in terms of computation in each node. However, there are two distinctive advantages in this network versus a conventional network, namely the availability of the nodes in large numbers and the flexibility of the deployment infrastructure. We argue that these two advantages not only play a significant role in reducing the energy consumption of each node, thereby extending the lifetime of such a network, but also provide us with a quality of sensing that is significantly better than that of traditional image sensors. As a result, in many situations it is prudent to use battery-operated image sensors instead of traditional sensor nodes. In this paper, we have presented a qualitative argument for using battery-operated image sensors. An analytical discussion will be the topic of our future work.

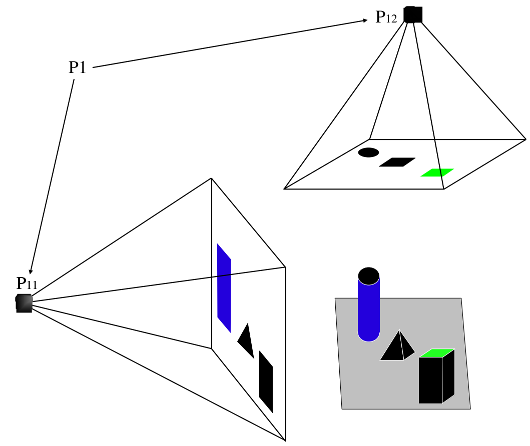

Figure 6. Distinct views of the objects reveal perspective-dependent properties of the objects.

6 References

[1] Phillip B. Gibbons, Brad Karp, Yan Ke, Suman Nath, Srinivasan Seshan, “IrisNet: An Architecture for a World-Wide Sensor Web,” IEEE Pervasive Computing, Volume 2, Number 4 (October-December 2003).

[2] Piyush Gupta and P. R. Kumar, “The Capacity of Wireless Networks,” IEEE Transactions on Information Theory, vol. IT-46, no. 2, pp. 388-404, March 2000

[3] Mohammad Rahimi ,Rick Baer ,Obimdinachi I. Iroezi , Juan C. Garcia ,Jay Warrior ,Deborah Estrin ,Mani Srivastava, “Cyclops: In Situ Image Sensing and Interpretation in Wireless Sensor Networks”, Third ACM Conference on Embedded Networked Sensor Systems (SenSys), November 2005.

[4] Shermer, Thomas, “Recent Results in Art Galleries,” Proceeding of the IEEE, Vol. 80, issue 9, September 1992.

[5] Brian Funt, Florian Ciurea, and John McCann “Retinex in Matlab,” Proceedings of the IS&T/SID Eighth Color Imaging Conference: Color Science, Systems and Applications, 2000, pp. 112-121.

[6] Martin Vetterli, “Fast 2-D Discrete Cosine Transform,” Proceedings of the ICASSP ’85, pp. 1538-1541.

[7] Sam Madden, Michael J. Franklin, Joseph M. Hellerstein, and Wei Hong, “TinyDB: An Acqusitional Query Processing System for Sensor Networks,” ACM TODS, 2005.