A Multi-Task Grocery Assist System for the Visually Impaired

Peter Zientara, Siddharth Advani, Nikhil Shukla, Ikenna Okafor, Kevin Irick, Jack Sampson, Suman Datta, and Vijaykrishnan Narayanan

I. Introduction

According to the World Health Organization, "285 million people are estimated to be visually impaired worldwide" [1]. Several technologies such as automatic text readers, Braille note makers, and navigation assist canes have been developed to assist the visually impaired. Concurrent advances in computer vision and hardware technologies provides opportunities for a visual-assist system that can be used in multiple contexts.

As part of the Visual Cortex on Silicon program, we have been developing interfaces, algorithms and hardware platforms to assist the visually impaired with a focus on grocery shopping. Grocery shopping is an essential activity in our daily lives that involves various interconnected activities. These include checking the pantry for current inventory, making a shopping list based on planned meals, getting to the store, and making opportunistic and impulsive purchases in response to signage at the store. Each one of these activities poses a significant challenge without visual cues. Consequently, our Third-Eye prototype enables a combination of hardware-software mechanisms to interpret these visual cues and communicate them to a visually impaired user as verbal or vibrational feedback.

A potpourri of different vision algorithms spanning brain-inspired algorithms, structured feature extraction techniques and deep learning approaches is used to support the automation of the different visual tasks. While both the origins and implementations of these algorithms are diverse, many share a common feature in that they work to reduce the potentially vast search spaces that vision problems can involve. Brain-inspired solutions, such as saliency algorithms, help to focus attention onto only specific parts of a complex image, thereby significantly reducing the computational effort to process the input and can act as a filter for subsequent steps. Other brain-inspired solutions such as GIST provide a contextual reference to the image and prime the decision making with that information, thereby filtering the model space of likely objects to be found.

In composition, these algorithms can be very powerful. Consider searching for an object in an aisle. First, steps such as saliency can help you reduce the effort spent examining the floor, empty shelves and other parts of an image that don't strongly register as potential grocery objects. Then, by understanding the context of specific grocery aisles instead of the entire store, complexity can be further pruned. For example, if milk is in your shopping list, the system can be primed to only search for milk when you arrive at the dairy aisle and to furthermore limit the number of distinct types of possible objects considered during classification to those likely to be in a dairy aisle. Such optimizations considerably reduce the computational load on the assistive vision system. In conjunction with these algorithmic advances, we have developed customized hardware solutions to make these operations more power-efficient as well as provide real-time feedback to the user.

The rest of this article describes the following set of visual assistive functions as shown in Figure 1: (1) Identifying objects in a pantry including misplaced items; (2) Identifying other shoppers when navigating in the store or when navigating to the stores; (3) Locating packaged objects from the grocery shelf and picking them; (4) Assisting in identifying items from prepared food sections.

For each of these tasks we consider not only the computer vision aspects of the problem, but also the practical embodiment of a complex system spanning wearables, edge, and cloud computing platforms that must solve the last mile issues of a real, usable consumer device.

II. Platform and Interfaces

One of the most important aspects of any technology is how the user interacts with it. Having an intuitive, simple, and functional interface can often be the difference between a successful widely adopted device and that which is not. The interface design becomes even more important when trying to help someone with an impairment. Not only does the interface have to be user-friendly, but it also has to be robust for different environments. In the case of assisting a person with visual impairment this means being able to handle cases which people without visual impairment handle without even realizing they do. Such a scenario would be when reaching for a product if you momentarily move your head in a different direction.

Our visual assistive system consists of several interacting components:

- An off-the-shelf Android-powered smart glass. The glass has both a camera and a built-in headset along with networking capability.
- We use a specially designed prototype glove that has been modified to have both a camera attached to it, as well a set of vibration motors.
- A shopping cart that can be equiped with a computer and a variety of sensors that would be provided by the retail location.

Fig. 1. Applications and underlying technologies for visual assist in grocery shopping.

- An IBM CAPI-enabled [2] high-performance server with custom tightly integrated FPGA acceleration.

Figure 2 shows a person wearing the smart glasses and gloves during a test of the system. The compute-augmented shopping cart (see 1) is not shown.

A. Interfaces

For our assistive technology we employ two main modes of providing feedback and guidance to the user. These modes are auditory feedback and tactile feedback. To provide this feedback to the users, we use the glove and the glasses as listed above.

1) Smart Glass: The off-the-shelf smart glass provides the system with a head view as well as network connectivity and speakers for audio feedback. In the assistive system the glasses are mainly used to guide the person at the aisle level to be in front of their intended/desired product. The commands such as "left", "right", "forward", "back" provide the necessary direction.

2) Custom Glove: The custom glove we employ has both a camera and a series of vibration motors. This camera that is on the glove allows the system to have the view point of what the person is reaching out for. This viewpoint may be different from that of the camera mounted on the headset and is critical to being able to provide guidance all the way to physically picking up the intended product. The attached vibration motors allow the glove the system to provide subtle feedback to the user to convey to them which direction they would have to move their hand to be able to grab the desired product. An example of this would be buzzing the right motor to indicate a rightward motion or the top motor to indicate the person needs to lift their hand.

Fig. 2. A person using an assistive system using multiple modes of feedback.

B. Using the System

While certain aspects of the system differ across the particular tasks it supports, the modes and mechanisms employed during grocery shopping provide good coverage of typical operations, and we describe them in detail below. In our system, the auditory feedback combined with the haptic feedback from the glove provide the needed assistance to the shopper.

C. Challenges

Creating a truly assistive system with a variety of interfaces presents a series of challenges, not all of which are initially obvious. These challenges include guiding the person through the store which includes the challenge of localization, obstacle and person avoidance, and grocery shopping. Other challenges are user centric. These include adapting the frequency of guidance commands to the speed at which the person is moving, reconciling different camera views to provide correct guidance, and having enough computational power to keep the system real-time. These challenges can be resolved via various methods. To solve the problem of guiding the person through the store, the smart cart could be equipped with various sensors. This could include cameras that not only have RGB information, but even depth and possibly thermal sensors. The use of localization technologies such as indoor GPS, and Bluetooth beacons around the store have the ability to track the user and provide the needed level of localization to the system (e.g aisle location). The challenge of reconciling different camera views arose from having two camera views that are not always in alignment with one another (the glove and the glasses). An example of when this occurs is while shopping when the user goes to grab a product, they might look away while still reaching in towards the intended product. This poses a challenge to a system giving guidance based on the view from those cameras. One possible solution to this issue would be the addition of sensors to the glasses and the glove. The addition of an IMU and Magnometer to both of the edge compute solutions give the ability to correctly provide guidance in this case, and other similar cases. An example how this would work is if the headset camera view indicates the person needed to move right, but the glove camera was pointing straight. The system would be able tell the user to turn just their head to align the two views, rather telling them to step right. However while all of the listed challeges can be considered implementation details, the biggest challenge that exists in an assistive system is being able to keep up with the real-time demands of the user. With an assistive system, solving challenge this is critical. In order to do this effectively, the system as a whole must leverage all available compute, including the compute power, however limited, available at the edge devices and the local infrastructure. As stated earlier, for our cloud compute device we use a high-performance server that is enabled with both FPGAs and GPUs. By leveraging custom architectures and exploiting parallel algorithms we are able to process 1080p video frames at around 50fps. While this may seem like it meets the real-time constraint it does not. This is because a server needs to be able handle multiple connections at once. Our current accelerated back-end would be able handle about 50 streams at 1fps. To make up this gap in performance, tricks need to be played at the local and edge compute devices to make this disparity become imperceivable. Some of the compute that can be offloaded the the edge devices and local infrastructure are mainly filtering proccesses. For instance, during the product detection phase a local infrastructure would be able to run the images being streamed back to the server through one of our hardware accelerated saliency algorithms. Additionally the edge device could use its sensors to only send a frame when the user has moved enough that the scene needs to be recomputed fully. Once the products are detected, the local infrastructure or edge device would be able to run a computationally less demanding tracking algorithm to be able to continue to guide the user toward the product between communications with the cloud back-end. In the next section we discuss a few algorithms deployed in our recognition system that use techniques to augment machine learning.

III. Beyond Deep Learning: Pragmatic Optimizations for Constrained Resources and Limited Time

Deep learning architectures like belief networks and neural networks are accelerating the pace of innovation in various industries such as autonomous systems, retail shopping and social media. These complex computational models are used to churn large amounts of data so as to make predictions on new observations. For example, consider an autonomous vehicle driving through a busy street. The decision of stopping the vehicle will depend upon whether there is an impeding obstacle or a red signal. More importantly this decision will be needed to made in a small amount of time and may be based upon multiple noisy sensory inputs.

Convolutional Neural Networks (CNNs) have become extremely popular and being used to solve a variety of image recognition and computer vision tasks. More recent and advanced CNN architectures have become deeper and more complex having 10 to 20 layers of Rectified Linear Units, hundreds of millions of weights, and billions of connections between units. The reader is pointed to [3] for insights on deep architectures in general and [4] for CNN-based learning and their recent advances.

While CNNs are an important thrust of research, they tend to be computationally expensive and deploying them on mobile platforms results in huge memory overheads. Another important insight recently unearthed by [5], is that the accuracies of CNNs can saturate after a few million images of training data. Also the overall efficacy of the image recognition pipeline is contingent upon having a good region proposal scheme that feeds regions of interest (RoIs) into the CNN. Considering these challenges, a variety of strategies can be used to augment the capabilities of neural networks and we outline a few below.

A. Visual Attention

Humans process and respond to only certain streams of visual information depending upon the task at hand. From a systems perspective, visual attention can be used as an efficient mechanism to prioritize data processing. Pixel-level saliency models such as Attention by Information Maximization (AIM) [6] can be used in automatic household pantry organization and maintenance, particularly as part of assist systems for the visually impaired. Computationally, AIM determines visual salience based on the amount of information present in local regions of the image within the context of its surrounding region. Suppose a product (say cookies) is wrongly placed in a shelf that stores products of another type (say shampoo), the segment of the shelf image containing cookies is “less likely” (higher self-information) to appear in the scene which mostly has image patches of shampoo, and therefore it is easily distinguishable or is considered “salient”. This is shown in Figure 3. Once a salient region is detected, a second stage of object classification can be deployed to identify the wrongly placed object.

(a) Original image (b) Thresholded saliency map

Fig. 3. Saliency used for missplaced item detection.

Figure 4 shows how a clustering of feature points can intelligently segment items in a grocery shelf. Our approach works well even with small and closely placed items of different types, as are commonly seen in grocery environments.

(a) Original image (b) Segmented heatmap

Fig. 4. SURF keypoint clustering used for product segmentation.

At the other extreme, while grocery environments are diverse, in a given grocery aisle there are often a lot of similar looking products like cereal boxes, detergent bottles, etc. Rather than processing each of these items as independent entities, we can localize similar RoIs, run our classification engine on only one of them and then assign the corresponding label to the entire group of similar RoIs. Figure 5 illustrates this flow where AIM is used to generate initial seed RoIs that are then coupled with Speeded Up Robust Features (SURF) keypoint matching to generate a list of RoIs that are similar in structure.

B. Context

While deep learning models use learnt features to recognize objects in a scene, visual systems can augment the current input data with knowledge about the usual organization of the environment that can provide greater scene context [7] and heirarchical models [8] that can reason about objects as collections of parts. Compositional rules can be used to build context cues to recognize objects never seen before if they contain previously seen sub-pieces. For example, if a product’s packaging changes or displays a seasonal advertisement, but key portions of the branding remain, then the product could still be recognized even if this version has never previously been seen. Similarly, when it comes to recognizing objects in video streams, spatial and temporal context can play a huge role in reducing the workload on computationally intensive models such as CNNs. For example, in [7], the authors proposed a Bayesian network called the Visual Co-occurrence Network (ViCoNet) where objects were represented as nodes and edges represented spatial relations between them. Figure 6 shows an example of a subgraph in such a system, which could very well represent a fresh-fruit section of a grocery shop. From this graph, it can be seen that, if the previous classification is a peach, then the next region is more likely a plum rather than mango, and this probability can be used to bias the classifications. Moreover, this graph naturally encodes aisle relationships, and can be used to feed an aisle predictor to narrow the set of plausible classes for any identification. These pruning features allowed the system using this network to improve performance as well as recognition rates.

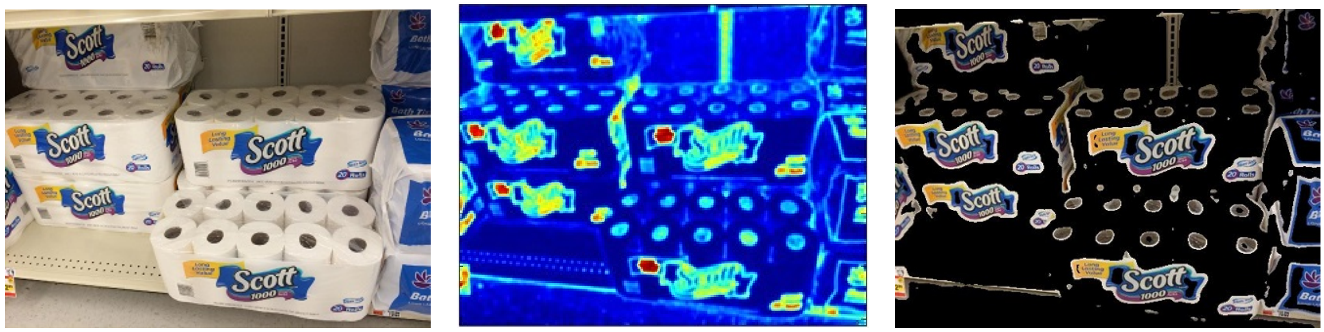
C. Multimodal Fusion

Humans use multisensory information from different sensory systems and combine it to influence perception, decisions, and overt behavior [9]. Wearables can be used in a similar fashion to help users in different tasks. A rich topic of exploration is figuring out a way to fuse multi-sensor information, especially data from vision that is fundamentally two-dimensional with a temporal unidimensional stream of data from other sensors to make predictions of the current state of the user. Multi-sensor information coming in from different devices can be streamed to distributed networks that can then make real-time updates. In Figure 7, we illustrate data recorded from a wearable device while two users walk in three different directions. As can be seen, these sensors are sensitive enough to be used as localization cues.

In the next section, we move further down the stack and discuss relevant system-level design and architectures that can further enhance the performance of mobile visual-assist systems.

IV. Hardware Support

Even with significant investments in algorithm development and selection, the computational costs of running a visual assist system on traditional computing platforms can still be significant and a potential impediment to practical deployment. In this section we describe efforts to implement custom designs for assistive vision, leverage new, non-traditional architectures, and forecast the impact of emerging technologies that can further improve the efficiency and effectiveness of visual assist systems from the wearable front-ends and mobile

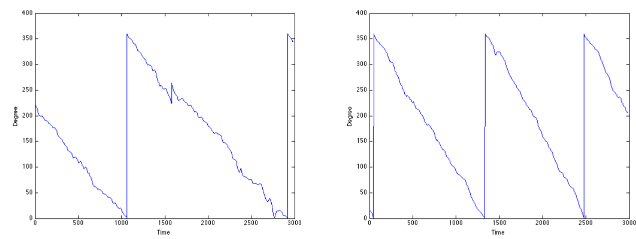


(a) Original image (b) Saliency Map (c) Thresholded Map

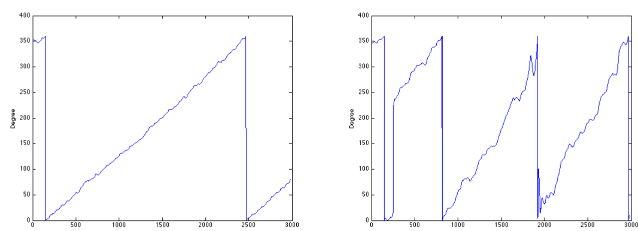
(a) Best RoI (b) SURF keypoints (c) Regions of Interest

Fig. 5. Saliency and SURF used to identify similar items.

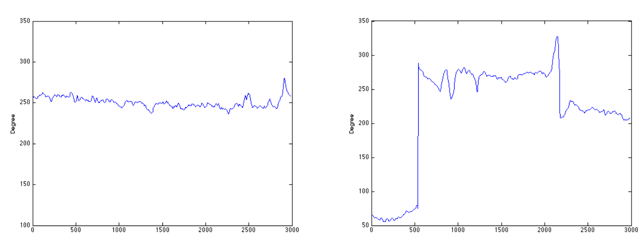
Fig. 6. Spatial relations exist between frequently co-occurring objects. These relationships can then be used as context cues to guide the recognition task.



(a) Person 1 - Left (b) Person 2 - Left



(a) Person 1 - Right (b) Person 2 - Right



(a) Person 1 - Straight (b) Person 2 - Straight

Fig. 7. Temporal sensor information used for localization.

edge computing platforms through the cloud-hosted back-ends and databases.

A. Custom Chips

While the computational needs of the computer vision algorithms we employ are substantial, they are also heavily structured and ammenable to acceleration. The computation and memory management for sub-tasks, such as person-detection or recognizing sets of replicated objects on grocery shelves, can be heavily customized to yield both large performance and energy gains. We have developed FPGA-based solutions for these sub-tasks, although our designs could also translate to ASIC implementations.

In [10], the authors proposed a scalable solution for object detection using structured features - Histogram of Oriented Gradients. While these features are very lightweight in terms of hardware resources, the high miss-rate is cause for concern. On the other hand, a deep CNN would have better accuracy, but would struggle to meet stringent power and area constraints when being deployed onto a wearable platform. Our current work is focussed on reducing the detection threshold and using these detection outputs as region proposals to a shallow CNN. We use three convolutional (+ pooling) layers and three fully connected layers in our CNN. The training images were of size 64×128 with around 3000 positives and 6000 negatives and we used stochastic gradient descent with a step-down approach to the learning rate. Figure 8(d) shows the output of our structured HOG custom hardware coupled with our shallow CNN.

B. Brain-like Architectures

The increasing prevalence of brain-inspired algorithms, such as saliency, and the dramatic rise in the use and utility of machine learning workloads has inspired research into architectures that more directly embody brain-like functionality. These new architectures, such as IBM's TrueNorth [12] completely abandon the traditional, centralized von-Neumann architectures of general purpose computing for distributed, neuro-inspired computation models. While moving tasks to these new brain-like architectures generally requires significant rethinking and re-expression of the existing codebases, for tasks that were already modeling neural networks, such as many machine learning workloads, the impacts can be both rapid and profound.

IBM's TrueNorth chip is an architypical example of this new class of brain-like architectures. It consists of 4096 neurosynaptic cores arranged in a 2-D array occupying 4.3 cm^2 of area in a 28 nm low power CMOS process. A key advantage of this chip is that it consumes merely 65 mW of power while running a typical computer vision application [12]. Having accelerated some of the key computer vision models using custom fabrics like FPGAs and GPUs, we are now looking to map them onto TrueNorth. While our FPGA and GPU acceleration efforts have supported real-time performance levels and greatly increased power efficiency over traditional software approaches, moving to these new brain-like architectures offers the potential to push computation even closer to the sensing platform by easily operating within wearable power budgets.

For example, extracting a HOG-like feature vector for a given 64×128 RoI would require 23 cores consuming around 62 mW of power. To evaluate the fidelity of this feature, we trained an SVM using the INRIA dataset [13] and ran evaluations. Figure 9 (b) depicts the dot product output between the trained model and the TrueNorth HOG feature model, while Figure 9 (c) shows the thresholded detections when evaluated on a test image. These brain-like architectures can eventually help move computation closer to the sensor when being used in constrained environments where cloud support may not be available. Also their low power consumption make them viable solution for camera-enabled wearable technologies.

C. Emerging Devices

In addition to advances in novel architectures for vision, there are also new emerging technologies on the horizon that can potentially offer new visual computing paradigms. One such promising technology involves harnessing the "analog" dynamics of coupled oscillators, which represents a radical departure from the conventional 1 and 0 based "digital" computing. Essentially, such oscillators are highly miniaturized, very low-power electrical devices, built using new functional materials such as vanadium dioxide (VO_2), that output a periodic electrical signal (voltage, current).

Coupled oscillator based computing entails the interconnection or coupling of a number of such oscillators; and the computation is performed in the way such highly-interconnected systems of oscillators synchronize or "talk" to each other. This alternate computing scheme is being particularly explored as a transformative solution for application in artificial vision systems that would be used in technologies of the future such as autonomous vehicles, augmented reality.

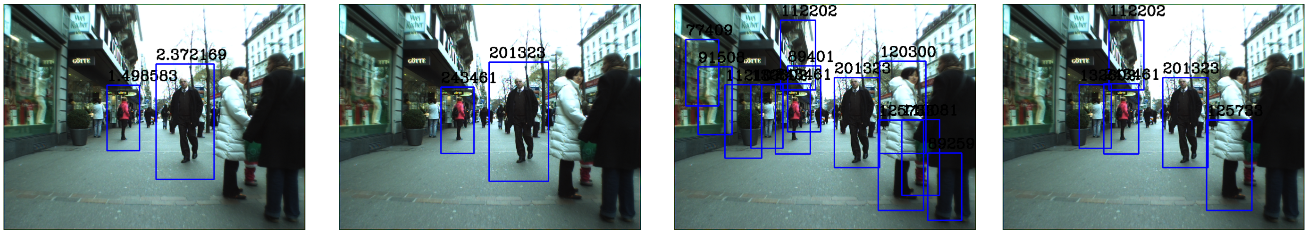
There is an existing body of work that shows how the analog functions of weakly-coupled oscillators can be used to provide distance-like metrics as a new analog primitive [14] and how sets of oscillators can implement analog convolutions and other useful high-level computational primitives. Recent advances in materials and device technologies offers the promise of nano-scale oscillators, such as hyperFET-based oscillators [15], that can implement these functions in extremely small area and power budgets.

Nano-oscillators are particularly intriguing for wearable vision applications. Early explorations indicate that arrays of hyperFET-based oscillators are small enough and consume sufficiently little power to be directly integrated into the image sensing chips. In addition to the benefits from offering higher-level analog primitives that improve computational efficiency relative over digital calculations, recent work has highlighted the large system-level benefits of moving early computation to the sensor chip [16], thereby limiting the losses incurred in moving data from the sensing to computing portions of wearable and mobile systems.

Fully exploiting the potential of these nano-oscillators will, however, require substantial remapping efforts for existing computations. Our group has mapped a number of image pre-processing primitives onto coupled-oscillator arrays as shown in Figure 10, and efforts are underway to map larger computations, such as HOG.

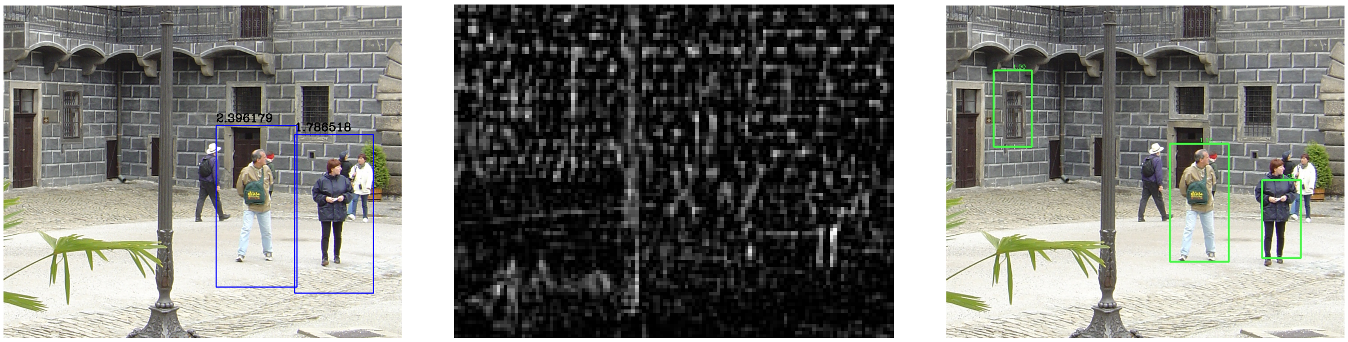
V. Conclusion

This work highlighted the efficacy of personal visual assist systems in our day to day activities. More specifically, the design and integration of a multi-task grocery assist system for the visually impaired was discussed. We also introduced the latest architectures and emerging devices that are being explored to further improve the capabilities of such systems. We continue to integrate text recognition to further improve product recognition; enable dynamic recommendations as we test our system in more challenging and distributed environments; and provide querying to personalize recommendations as part of a holistic solution for the visually impaired.



(a) OpenCV output (b) Hardware HOG output [10] (c) Reduced threshold (d) HOG-CNN output

Fig. 8. Coupling structured features with learned features. Input image obtained from [11].



(a) OpenCV output (b) TrueNorth Dot Product (c) TrueNorth Detections

Fig. 9. Mapping HOG to True North. Input image obtained from [13].

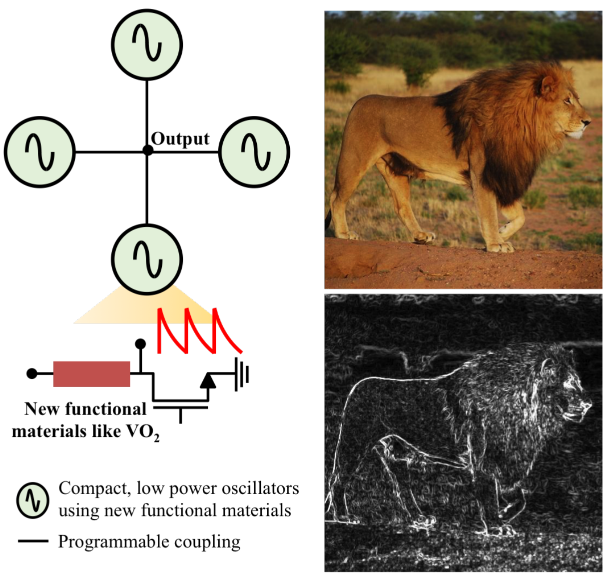


Fig. 10. Edge detection using coupled oscillators

VI.

Acknowledgment

This work was supported in part by NSF Expeditions in Computing Program: Visual Cortex on Silicon CCF 1317560. We acknowledge all collaborators and participants of the Center whose work has influenced this article.

References

[1] World health organization. (Date last accessed 16-Sep-2016). [Online]. Available: http://www.who.int/mediacentre/factsheets/fs282/en/

[2] J. Stuecheli, B. Blaner, C. R. Johns, and M. S. Siegel, "Capi: A coherent accelerator processor interface," *IBM Journal of Research and Development*, vol. 59, no. 1, pp. 7:1–7:7, Jan 2015.

[3] Y. Bengio, "Learning Deep Architectures for AI." Now Publishers, 2009.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," pp. 436–444, May 2015.

[5] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," *CoRR*, vol. abs/1503.03832, 2015. [Online]. Available: http://arxiv.org/abs/1503.03832

[6] N. Bruce, "An Information Theoretic Model of Saliency and Visual Search."

[7] S. Advani *et al.*, "Visual co-occurrence network: Using context for large-scale object recognition in retail," in *Embedded Systems For Real-time Multimedia (ESTIMedia), 2015 13th IEEE Symposium on*, Oct 2015, pp. 1–10.

[8] I. Kokkinos and A. Yuille, "HOP: Hierarchical Object Parsing," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, June 2009, pp. 802–809.

[9] B. E. Stein, T. R. Stanford, and B. A. Rowland, "The Neural Basis of Multisensory Integration in the Midbrain: its Organization and Maturation," *Hearing Research*, vol. 258, no. 1, pp. 4–15, 2009.

[10] S. Advani, Y. Tanabe, K. Irick, J. Sampson, and V. Narayanan, "A scalable architecture for multi-class visual object detection," in *Field Programmable Logic and Applications (FPL), 2015 25th International Conference on*, Sept 2015, pp. 1–8.

[11] Robust Multi-Person Tracking from Mobile Platforms. (Date last accessed 16-Sep-2016). [Online]. Available: https://data.vision.ee.ethz.ch/cvl/aess/dataset/

[12] F. Akopyan *et al.*, "Truenorth: Design and tool flow of a 65 mw 1 million neuron programmable neurosynaptic chip," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 10, pp. 1537–1557, Oct 2015.

[13] INRIA person dataset. (Date last accessed 16-Sep-2016). [Online]. Available: http://pascal.inrialpes.fr/data/human/

[14] J. A. Carpenter, Y. Fang, C. N. Gnegy, D. M. Chiarulli, and S. P. Levitan, "An image processing pipeline using coupled oscillators," in *2014 14th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA)*, July 2014, pp. 1–2.

[15] W. Y. Tsai *et al.*, "Enabling new computation paradigms with hyperfet - an emerging device," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 2, no. 1, pp. 30–48, Jan 2016.

[16] R. LiKamWa, Y. Hou, J. Gao, M. Polansky, and L. Zhong, "Redeye: Analog convnet image sensor architecture for continuous mobile vision," in *International Symposium on Computer Architecture*, 2016.

Peter A. Zientara (paz117@cse.psu.edu) received the B.S. degree in computer engineering from the York College of Pennsylvania in 2014 and is currently pursuing his Ph.D. in computer science and engineering from the the Pennsylvania State University. His research interests are in smart sensors, hardware acceleration, and embedded vision systems.

Siddharth Advani (ska130@cse.psu.edu) received the B.S. degree in electronics engineering from the Pune University, India in 2005, the M.S. degree in electrical engineering from the Pennsylvania State University in 2009 and the Ph.D. in computer science and engineering from the Pennsylvania State University in 2016. His research interests include embedded vision design for smart mobile applications targetted for domain-specific applications.

Nikhil Shukla (nshukla@nd.edu) received the B.S. degree in electronics and telecommunications engineering from the University of Mumbai, India in 2010. He is currently working toward the Ph.D. degree in electrical engineering at the University of Notre Dame.

Ikenna Okafor (izo5011@cse.psu.edu) is currently in the Integrated Undergraduate program for computer engineering at the Pennsylvania State University. His research interests include hardware acceleration for neuromorphic vision algorithms.

Kevin Irick (kevin.irick@siliconscapes.net) is the CEO and Founder of SiliconScapes LLC based in State College, Pennsylvania. Kevin founded SiliconScapes with the objective to develop high-performance, embedded and cloud-based, video analytics systems for security, surveillance and retail applications. Kevin conferred the Doctoral degree in computer science and engineering from the Pennsylvania State University in 2009. His research activities included application-specific hardware accelerator design methodology, hardware-assisted image processing and recognition, and high performance computing on reconfigurable architectures.

John (Jack) Sampson (sampson@cse.psu.edu) received the B.S. degree in electrical engineering and computer science from the University of California, Berkeley in 2002 and the Ph.D. in computer engineering from the University of California, San Diego in 2010. He is an currently an Assistant Professor with the Department of Computer Science and Engineering at the Pennsylvania State University. His research interests include energy-efficient computing, architectural adaptations to exploit emerging technologies, and mitigating the impact of Dark Silicon.

Suman Datta (Suman.Datta.1@nd.edu) received the B.S. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1995, and the Ph.D. degree in electrical and computer engineering from the University of Cincinnati, Cincinnati, OH, USA, in 1999. From 1999 to 2007, he was a member of the Logic Technology Development Group, Intel Corporation. He is currently a Professor of Electrical Engineering at the University of Notre Dame and his group is exploring new materials and novel device architecture for CMOS enhancement and replacement for future energy-efficient computing applications.

Vijaykrishnan Narayanan (vijay@cse.psu.edu) received the B.S. degree in computer science and Engineering from the University of Madras, India, in 1993 and the Ph.D. in computer science and engineering from the University of South Florida, Tampa, USA, in 1998. He is currently a Professor of Computer Science & Engineering and Electrical Engineering at the Pennsylvania State University. His research interests include power-aware and reliable systems, embedded systems, nanoscale devices and interactions with system architectures, reconfigurable systems, computer architectures, network-on-chips, domain-specific computing.