WAREHOUSE MOBILITY BEYOND VOICE AND RF

Emerging smartglass and augmented reality technology has the potential to transform picking and other DC applications over the next decade. This next phase of mobile transformation continues a trend that started in the early 2000s with the introduction of multi-modal mobile solutions that merged the best of traditional RF/barcode scanning systems and voice-only picking applications. Today's smart mobile applications combine voice direction, speech recognition, device displays, and barcode scanning and run on powerful, lightweight, and cost-effective Android handheld and wearable devices.

Today's smart mobile applications enable process and workflow innovations that deliver levels of accuracy and productivity far beyond what was possible with earlier voice-only or scan-only systems. In the near future, these applications will run on smartglasses and generate additional efficiencies by transforming how users receive and interact with visual information.

This paper outlines the long-term technical and financial forces that are driving the evolution of mobile applications in the DC. We start by describing the transition from voice-only warehouse applications to today's voice-scan-and-display based apps that run on smartphones and new warehouse wearables. In the final section we look ahead and explore how maturing RFID, smartglass, augmented reality, and vision recognition technology offer the promise to further transform manual processes, making work easier, faster and better for DC associates.

First Generation Voice: Single-Purpose Hardware

In the late 1990s voice-directed warehouse applications emerged as a proven solution for improving distribution efficiency in industries ranging from grocery and foodservice distribution to retail and medical supply. The first-generation voice applications ran on rugged, mobile computers and combined a voice-directed workflow (the system provides audio prompts directing users what to do) with speech recognition technology that understands a user's spoken responses.

First-generation voice-directed applications were delivered on wireless, voice-only hardware appliances that had a headset connector and a few buttons, but no display, no keypad and no scanner. They were designed and manufactured by voice hardware specialists that closely tied their own speech recognition technology to their single-purpose hardware platform. Some of these early voice appliances included Digital Signal Processor (DSP) chips and other specialized circuitry to execute the speech recognition algorithms.

In 1998, the founders of Lucas Systems set out to deliver hardware-independent voice applications for the warehouse that could run on open, industry standard computers that included keypads, touch screen displays, and barcode scanners. At that time, the audio circuitry of standard mobile computers used in a distribution center did not provide the high quality, high resolution audio processing required

The early voice appliances made voice in the warehouse a reality, but the downside was that the hardware was expensive and available only from a few small companies.

for acceptable speech recognition accuracy. As a result, Lucas delivered its warehouse voice applications on the voice appliances manufactured by voice-only hardware companies, including Voxware[1] and Vocollect.

The downside to proprietary devices is that the hardware and accessories are expensive and only available from a few companies, leading to a high degree of vendor lock-in. The founders of Lucas anticipated that general-use mobile computers would supersede proprietary voice-only appliances, similar to the history of every other computer market (see Are Voice-only Terminals Headed for the Trash Heap of Obsolete Technology?)

Second Generation Software and the Transition to Multi-Modal Hardware

In the early 2000s major consumer products companies, retailers, and other distributors joined Lucas in pushing for alternatives to the special-purpose appliances available from the niche hardware companies. Voice users were motivated by risk avoidance (i.e., they didn't want to be dependent on small, single-product hardware companies), cost reduction, and the desire for more and better software options.

The major mobile computer makers responded. Lucas Systems worked closely with Symbol Technologies (now Zebra technologies) and several early adopters to develop a voice-capable model of Symbol's flagship mobile computer, the MC9060[2]. Early in 2005 the first Lucas voice applications were delivered using open, multi-modal hardware and industry-leading speech recognition software, proving the capability of this new approach. (Speech recognition technology had been undergoing a similar transformation as voice hardware – this is the subject of a separate Thought Leadership Paper: The Evolution of Speech Recognition Technology For the Warehouse[3]).

Since 2005, the audio and other capabilities of the industrial devices made by Zebra and other manufacturers have advanced. Today, Zebra offers a full range of voice-capable devices ranging from the lightweight and full-featured TC21 to the wearable RS5100 and the ultra-rugged MC33 (the successor to the MC9000 family that was used in the first multi-modal voice solutions). All other hardware vendors have followed suit with more powerful devices designed to support voice, giving DCs more options than ever.

Consequently, the market for single-use voice appliances has been shrinking. In 2005 Voxware announced plans to exit the hardware business and licensed its proprietary design to LXE. In early 2011, Vocollect, the only remaining independent voice hardware manufacturer in the US, was acquired by Intermec[4], and Intermec (and Vocollect) was subsequently acquired by Honeywell (which earlier acquired LXE).

While lower costs and greater versatility were the major factors that initially drove DCs to adopt multi-modal computers, the introduction of new voice applications that used voice alongside scanning and touchscreen displays accelerated the transition to multi-modal hardware. Even before the transition to multi-modal hardware began, Lucas recognized the efficiency advantages of combining voice

with scanning in a voice-directed workflow – our first voice and scanning solutions were deployed in 2003. While other companies that had their roots in the development of voice-only hardware continue to focus on voice-only software, the very first generation of Lucas software was designed for multi-modal input: it doesn't matter whether input is spoken, scanned, or entered on a keypad or touchscreen.

Starting in 2005, the market moved to the use of voice-capable RF devices such as the Zebra MC3000.

The Next Generation: Rugged Smartphones and New Wearables

In parallel to the evolution of the industrial voice hardware market, the introduction of the iPhone in 2007 started a revolution in the consumer mobile market. Smartphones represent a mass market for phones that double as mobile computers that are designed for applications combining display and touch entry, imaging, voice direction, and speech recognition, similar to the smart mobile apps that have emerged in the DC.

Smartphones and new consumer tablets have invaded the enterprise, initially in the hands of employees bringing their own devices to work. The BYOD trend created new security and device management challenges for corporate IT departments, while simultaneously opening the floodgates to new mobile applications in the enterprise. Today smartphones and tablets are an integral part of enterprise computing with well-defined security and management products and technology.

Beginning in 2014, Lucas offered customers the opportunity to run its applications on Android smartphones. Additionally, Zebra and other industry hardware manufacturers have adopted Android for the enterprise market. They now offer a range of Android touch computers that integrate scanning in highly competitive form-factors and price-points.

Smart Wearables

Beyond traditional mobile devices, a new wave of smartwatches and smartglasses have entered the market, opening new application possibilities in the warehouse. Besides a smaller, lighter form-factor, these devices make it easier to present visual information within a voice-directed workflow. For example, Lucas Move applications include display information and user help that are presented on the mobile computer screen, but users can now receive this information – product images, lists of pick locations and quantities, etc. – on the screen of a smartwatch.

Smartglasses open up all-new possibilities to incorporate visual information in a user workflow. Several manufacturers already offer smartglasses providing a heads-up display (HUD) – essentially a micro display device positioned on an eyeglass frame. The microscreen provides a viewable area roughly equivalent to that of a large smartphone or small tablet, and the devices include a speaker

for voice direction, a camera which can be used to capture still or video images and/or read barcodes, and a microphone to permit the use of speech recognition.

Next-generation augmented reality smartglasses will project images on a transparent lens mounted in an eyeglass frame, providing an augmented view of the operating environment, including visual cues and indicators in specific locations within the user's field of vision. This technology will fundamentally transform how visual information is presented and used by workers. While HUD and AR smartglasses still face some technical obstacles before they can be adopted for general use in the DC, those obstacles are falling. (see The Hardware Buyers Guide)

Choosing the Right Tool for the Job

The flexibility of today's smart mobile applications makes it important that engineering and operations teams closely consider when it is best to use voice, when to scan, and when to use a device display.

When to Scan

Almost any voice-directed warehouse process could use speech recognition or barcode scanning to confirm tasks and capture information. For example, in a basic picking process, a user can speak a check string printed on a shelf location to confirm he is picking from the right location, or he can scan a barcode label on the location. In this example, speaking is faster than scanning, even if the user has a finger-mounted ring scanner. With voice, the user will speak the check string while approaching the slot and reaching for the product. With scanning, the user must stop, aim, and scan before reaching for the items.

The time penalty for using a scanner is small, but in high volume pick operations every extra second per transaction adds up to a significant number of man-hours over days, weeks and months.

In terms of accuracy, speaking a check string is as accurate as scanning a location label[5], but voice offers the advantage of requiring the picker to specifically confirm the number of

items he picks, reducing pick quantity errors. And to eliminate mistakes due to confusion over unit of measure or pack factors, users can be required to speak the pack type (each, bottle, case, inner pack, etc.) of items as they are picked.

On the other hand, for companies that need to capture additional product data at the point of pick, barcode scanning can be a better alternative than voice – both in terms of accuracy and efficiency. For example, it is faster to scan a serial number than to speak 8-12 digits. Moreover, the requirements to capture, track and trace product level data (lot, date code, etc.) are increasing in healthcare, food, and a wide range of other industries. In such cases, the ability to capture data at the point of pick offers big cost savings over post-pick data entry, adding to the financial advantages of voice-plus software. Depending on the frequency and type of data capture required, users can rely on the internal scanners that come with their

While other companies that had their roots in the development of voice-only hardware continued to focus on voice-only software, the very first generation of Lucas software was designed for multi-modal input: it doesn't matter to the application software whether the input is spoken, scanned, or keyed in.

devices, and in other cases ring scanners, and other small, wearable scanners can be a good complement to voice.

Another place where scanning can improve the efficiency of a voice-directed application is cart set-up in a batch picking process (i.e., where multiple customer orders are picked to separate totes or cartons on a cart). Using voice-direction and barcode scanning, the user can scan the barcode on a tote or carton as he places it on a picking cart, scanning the cart location to associate the carton ID with the cart location prior to picking. The voice-plus-scan approach to cart set up is much faster than either a voice-only or scan-only process, making this a great example of how voice and scanning together are better than either one individually.

Scanning is also a good fit in returns, receiving, put-to-store, and other processes in which entry of a product identifier initiates a task. For example, in certain put-to-store applications a user can voice or scan a carton, pallet or product barcode to identify the product and initiate the next step in the job. After scanning the barcode, the voice system tells the user how many of the item to pick and put into each tote, pallet or other container.

The user confirms by voice the put locations and quantities. In a voice-only world, the user could speak the item, SKU or other product number, but in some cases scanning can be faster than speaking. Since scanning in this example is used to initiate the task, there is no time penalty for handling a scanner.

In the end, whether voice or scanning is better in a given DC or process depends on a number of factors, including the availability of barcodes, the ergonomics of scanning (the ease of scanning with a ring scanner or with the built-in device scanner), the specific data capture needs, and other details of the process.

Data Capture Beyond the Barcode

Another major frontier in mobile application functionality is the addition of RFID and NFC (near-field communication). NFC is widely adopted in consumer payment applications and RFID continues a slow but steady adoption in retail and other industries. It is not yet widely used in the DC, but there is every reason to expect that wil change.

One scenario for RFID/NFC use in a voice-directed application is for location awareness. In this case an RFID reader would read a tag on a product or location as the worker reaches for product. The reader could be

embedded in a bracelet, glove, or other wearable device. While a barcode ring scanner still requires a user to take his hand away from the task to focus the scanner on the item barcode, RFID is a true hands-free technology that allows rapid verification, rich serial data capture (GTIN, lot, batch, catchweight, expiration date, etc.) and additional tracking capabilities (temperature, etc.). Compared to voice, RFID reduces the potential for human error (when a picker reads a location check string but picks from an adjacent slot). Plus, the use of RFID enables the simultaneous capture and confirmation of location, item and additional item-level data (serial number, lot, weight, etc.) eliminating the secondary data capture step in a voice-plus-scan process.

Display Information

Device displays are also widely used today in voice-directed warehouse applications, especially outside of the main voice-directed task. For example, logging on to the system using display and keypad is as simple as logging in to an ATM or Website. This is far easier and intuitive for first-time users than a voice-only process, in addition to providing greater security than voice-based log-in. Likewise, new user training is improved by combining verbal and visual information on the device display about how to use the voice system. By contrast, DCs with voice-only wearables typically need to purchase additional hardware devices for training purposes. Finally, the device screen can also be used to display the dialogue - prompts and spoken replies - within a voice application, which is a helpful tool in training and problem resolution.

New user training is improved by combining verbal and visual information on the device display about how to use the voice system.

Other uses for the device display include showing product images, product details, or lists of information – all of which are more appropriate for display than delivering the information verbally. For example, if a selector is confused about the pack factor when picking an item (what does one “each” or box look like?), the application can display an image of the item or package to pick. The device screen can also be used to display a list of base items in a picking assignment so a selector can pre-plan his pallet. The application could deliver this information verbally, but the display screen allows the user to see more information at once, which is more conducive to managing or making decisions.

Within the next five years, smartglasses will begin to transform the visual interface for users. In addition to providing text and images similar to what can be delivered today on a smartwatch or other device screen, augmented reality glasses will provide visual cues and indicators within the users line of sight. For example, as a user approaches a pick face, the picker might see a virtual “light” on the required pick slot. Likewise, with the addition of vision-recognition technology, the smartglass camera could provide automatic task confirmation based on user activity – the system would “see” that the user is placing items in the correct tote or provide a warning if the picker places an item into the wrong tote.

Conclusion

The transition from voice-only to full-featured multi-modal applications has accelerated as the number of multi-modal hardware options increase, hardware costs continue to decline, and innovative new wearable devices emerge. While it will be some time before heads-up displays and ubiquitous, reliable and cost-effective RFID readers enter the mainstream of DCs, we can expect innovative combinations of voice with scanning and screens to proliferate. These process and workflow innovations are driving new levels of accuracy and productivity above and beyond what was possible with earlier voice-only or scan-only systems. And the cost for using smart mobile applicationss continues to drop, further growing the market for these applications within and beyond the four walls of the warehouse.

About Lucas Systems, Inc.

Lucas Systems helps companies transform their distribution center operations and continuously adapt to changing market dynamics. We dramatically increase worker productivity, operational agility and accuracy and reduce the need for labor.

Lucas solutions are built on 23-plus years of deep process expertise and smart software using AI-based optimization technologies. Our solutions features Jennifer™, the brain, voice and orchestration engine that drives performance improvement gains. We help you make the smartest moves at the lowest costs with Jennifer™.

References

1 Some of the first Lucas software applications were delivered on hardware from Verbex, which was acquired and rebranded by Voxware in 1999. Voxware Acquires Verbex Voice Systems, Enters Industrial Speech Recognition Marketplace, TMCnet.com, February 5, 1999. http://www.tmcnet.com/tmcnet/newscti/cti1000157.htm

2 Symbol Technologies Launches Speech-Recognition Mobile Computing Solution for Warehouse Environments, August 2, 2005, http://www.symbol.com/category.php?fileName=PR-859_MC9060_Launch.xml

3 Intermec Completes Acquisition of Vocollect, March 3, 2011, http://www.businesswire.com/news/home/20110303006783/en/Intermec-Completes-Acquisition-Vocollect

4 First-generation voice picking systems used fixed check strings at every location as the primary confirmation method, making it possible for users to memorise the location numbers, impacting pick accuracy. Modern voice systems address the issue of memorisation through subtle process changes (periodic item-level UPC checks, for instance) and the use of variable check digits. With variable check digits, managers can change the sequence of digits that will be used on a daily or weekly basis, so users cannot easily memorise the code. For example, on Monday, users will use the first three digits of a five-digit check string, and later that week they may be required to use the final or middle three digits of the string.

5 Witold Bahr and Ming K. Lim, RFID Adoption Issues in the Warehouse, 19th International Conference on Material Handling, Constructions and Logistics, 15-16 October 2009 at: http://eprints.aston.ac.uk/15222/.