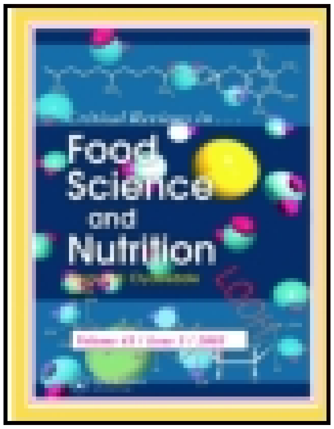

Critical Reviews in Food Science and Nutrition
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/bfsn20

An Overview of the State of the Art of Automated Capture of Dietary Intake Information

Robert Steele[a]

[a] Discipline of Health InformaticsThe University of Sydney, Sydney, NSW, 2000, Australia +61 2 9036 7368,

Accepted author version posted online: 20 Jun 2014.

An Overview of the State of the Art of Automated Capture of Dietary Intake Information

ROBERT STEELE

Discipline of Health Informatics

The University of Sydney

Sydney, NSW 2000, Australia

+61 2 9036 7368

robert.steele@sydney.edu.au

ABSTRACT

Significant benefits arise from being able to capture dietary or nutritional intake information automatically or semi-automatically. These include the ability for individuals to know and understand their nutritional intake and hence improve their diet and health. To date, only highly manual processes such as 24 hour recall, food diaries and food journals have been utilized which have been overly cumbersome for widespread adoption. Emerging informatics, computer vision, mobile computing and sensor-based approaches are likely to play a role in further automating the capture of dietary intake information and these are becoming increasingly utilizable through such advents as the rapid and ubiquitous uptake of smartphones with built-in digital cameras and other sensors. In this paper we review the state of the art of technologies for automatic capture of

dietary intake information and identify significant outstanding research problems and promising directions.

Keywords

dietary intake, automated capture, technology, informatics, nutrition, UPC, mobile health

INTRODUCTION

While there are significant benefits arising from individuals being able to capture their detailed nutritional intake information, doing so has remained a practically cumbersome, overly complex and manual task to date. Food journal or diary approaches exist, but they are largely based on human recall and/or manual input by an individual. This level of effort has proved prohibitive to their mass adoption and usage. An under-reporting bias present in manual approaches has also been observed (Martin et al., 2009a). At the same time, there are a number of informatics, computer vision, mobile computing and sensor-based technologies and approaches that can be applied to the capture of dietary intake information. The rapid and ubiquitous uptake of smartphones which now as a norm contain built-in digital cameras and a variety of other sensors provide a significant new person-centric data capture capability. In addition, there is the increasing development and prevalence of nutrition facts panel labeling (Nestle, 2010) and further digitization in food industry production and point-of-sale systems (Donnelly et al. 2012).

An implicit focus of a large body of work in relation to technologies for automating dietary intake capture has been upon dietary intake information capture in research or trial settings (Ngo et al., 2009; Illner et al., 2012). This recognizes the difficulties of automated dietary intake information capture even in such settings. In this paper we overview numerous technologies, and do not limit consideration of their applicability to just research trial settings – but in general provide some discussion of their practical applicability for non-trial and general community adoption and use.

It is posited that informatics, computer vision, mobile computing and sensor-based approaches will be important mechanisms to be further utilized in moving dietary intake capture to a more automated or semi-automated process. In this paper, we review the current state of the art in relation to such dietary intake capture technologies and identify outstanding topics of research and promising directions.

THE CASE FOR NUTRITIONAL INTAKE SELF KNOLWEDGE

The vast majority of the population do not know their aggregate nutritional intake for any given day or for longer periods, or even know their aggregate energy intake. This is all the more surprising given the compelling reasons for individuals to have self-knowledge of their nutrition: a healthy diet is a universal and daily issue for all members of the global population; the strong link between nutrition and health (Shils et al., 2006); the important role of nutrition in preventing and managing chronic conditions (WHO, 2003); the link between nutrition and longevity (Finch, 2007); the impact on economic development (Strauss and Thomas, 1998) and the large financial implications globally for the health system (Sturm, 2002). Notwithstanding recent advents such as more detailed food nutrition labeling adopted in various countries (Blumenthal and Volpp, 2010) and further diet-related public health interventions (Maruthur et al., 2009) the area of detailed self-knowledge of nutritional intake remains relatively and surprisingly under-developed.

While development of tables of component nutritional information per food has been ongoing for over a century, there are a number of recent developments, such as large digitally-stored nutrient databases, that provide significant further opportunities. A key technological development in recent years is the rapid uptake of powerful mobile or 'smartphone' devices that have various person-centric capabilities relevant to food consumption information capture and dietary decision making and guidance, such as advanced data capture, imaging and sensor capabilities, wireless and always-on connectivity, powerful processing and storage capabilities and high user engagement (Mutchler et al., 2011; Klasnja, 2009). Such devices also represent a data integration infrastructure that is being rapidly deployed throughout everyday life settings, does not require a government or employer to fund or install, and can be utilized for new forms of data integration to support health outcomes (Steele and Clarke, 2012).

Why pursue an informatics-based approach for individuals to know their nutritional intake in detail? As described above, contemporary technological developments now have the potential to make such capabilities achievable for mass use for the first time. In addition the capabilities provided by smartphones/ mobile devices appear to be superior to existing manual capabilities in various ways. Firstly, they will potentially enable an individual to understand and react to their full, detailed aggregate nutritional intake, which is not provided by individual food labels. Secondly, the digital capture of nutritional information allows the utilization of more advanced, context-sensitive computational aids to healthy eating, including simple display, context-sensitive recommendation systems and persuasive systems (Mutchler et al., 2011; Chatterjee and Price, 2009; Fogg, 2003; Deterding et al., 2011). Thirdly, the computational capabilities of mobile devices can be used to handle the full complexity of the detailed aggregate nutritional information – a complexity that prohibits the manual or paper-based recording and understanding of an individual's daily nutrition. Fourthly, such systems also allow the revolutionizing of current food labeling approaches – instead of a single static, physical nutrition label, a personally customized digital label can be displayed to individuals on their mobile device screen (or via other display) and they can select to look into more or less detail about a food's nutrition. Fifthly, there is the potential to interface such nutrition systems with health information systems such as Personal Health Records (Steele et al., 2012). Finally, personal nutritional knowledge gained through such systems can be combined powerfully with public health campaigns – as individuals are made aware of recommended dietary practices and policies (Hawkes, 2012), they can more easily check to what extent they are following these.

There is substantial evidence that even simple food labeling can affect healthier food choice (Barreiro-Hurle et al., 2010), that maintaining detailed records is a key technique for behavior change (Shay et al., 2006) and the benefits of dietary self-monitoring and recording to achieve weight-loss are widely demonstrated (Burke et al., 2011). However, further research into how the use of specific information technologies can affect health behavior change (Lindqvist et al., 2012; Jordan et al., 2011) is still required.

It could also be argued that availability of such information to individuals can empower their decision making and also potentially make organizations involved in food production, delivery and policy more accountable (Dutton and Eynon, 2009). A nutrition informatics infrastructure by making detailed nutritional information available to individuals could provide a platform for bottom-up or non-institutional innovation (Dutton and Eynon, 2009). Such information availability also promises to enable new forms of participatory research or 'citizen scientist' contributions.

AUTOMATED DIETARY INTAKE INFORMATION CAPTURE

ACCEPTED MANUSCRIPT

For an individual to know their dietary and nutritional intake, three broad approaches are possible. Firstly, for each meal or food item consumed, two pieces of information would need to be captured 1) the exact types of food eaten, each of which may or may not have an associated unique identifier and 2) the portion-sizes eaten. This information would need to be captured across all eating activity to gain a complete and exact view. For the case of capturing this information for the purpose of research trials, the period of capture is limited and consenting participants are involved – allowing a significant relaxing of the need for highly automated approaches.

Secondly, aggregate approaches are considered. While the foods and their quantities eaten at each sitting may not always be known, nevertheless a measure of total intake over a period can be formed. This approach recognizes that while not having exact capture per meal or sitting, aggregate approaches still provide valuable information in many cases. This approach is discussed below.

Thirdly, approaches based on estimation and extrapolation from past eating behavior. This approach may also result in a less exact capture of information but can nevertheless still provide additional and valuable information. This approach is also discussed below.

It should also be noted initially that computer vision, mobile computing and sensor-based approaches are not the only possible technology types to capture the needed information. But particularly if detailed information is needed about actual food consumption activities, then these technological approaches, given their interfacing with the 'real world' in terms of vision and physical sensing, are applicable to many aspects of this data capture, and may also be the only mechanisms available for some aspects of such information capture at this current time.

Type of Data Capture

In this paper we overview in turn, the automated capture of a number of different types of data that have a part to play in comprehensive, digital dietary and nutritional intake information capture. These types are as follows:

- Food type, where no food identifier available
- Food portion size
- Food identifier information: food identifiers including barcodes, are not always available, but when they are, they provide a powerful and unambiguous information input
- Nutrition facts panels
- Recipe information
- Aggregate intake-related information: particularly focusing on itemized food purchase/ point-of-sale data
- Estimation approaches

The third, fourth, fifth and sixth of the above list require the capture of structured information or text and so do not depend in all cases on advanced computer vision, image processing or sensor-based approaches, but at the same time such technologies will still help in many cases in relation to these data types.

ACCEPTED MANUSCRIPT

In the subsequent sections of this paper, capture of each of these types of data relevant to dietary and nutritional intake information will be considered in turn.

Time of Capture

While it might appear that time of information capture should be at the time of consumption this is not necessarily the case. For example capture at the time of purchase offers a number of benefits.

For example in the case of capturing nutritional information in aggregate it may be useful to simply capture a full grocery purchase list at time of purchase, also thereby allowing the calculation of the aggregate detailed nutritional information of the full grocery load (Brinkerhoff et al., 2011). This approach of course does not capture the specific foods or their portions that an individual actually eats at each sitting, and requires extra information capture to identify what foods or portions are eaten by different members of a multi-person household. This case will be discussed further under the section on aggregate approaches.

For completeness, the times of capture for information relevant to dietary intake: prior to purchase, at time of purchase, after purchase, at time of consumption or after consumption. For example, traditional manual 24 hour recall-based survey approaches (Karvetti and Knuts, 1985), and most manual capture approaches used in research trial scenarios, ask individuals to recall and record their dietary intake after they have eaten.

Practicality of Use

For each of the types of data being considered the practicality of capture will also be briefly discussed. This is because while the presence of smartphone-based digital cameras and sensors has a profound impact upon the mass population's ability to capture data, there are still barriers in terms of how consumers may or will use these devices. For example, it could be easier in some cases to utilize wireless transmission or data integration of point-of-sale purchase information, as such may require even less manual effort by individuals. Even what are considered highly automated approaches, such as those involving image analysis of food images captured by individuals photographing meals via a camera or smartphone (Zhu et al., 2010; Martin et al., 2009b), represent usage barriers for universal adoption and ongoing utilization, given that they require an individual to take such an explicit action as photographing.

At the same time, while practicality of use is an important aspect to consider, discussion of detailed evidence or analysis of comparative ease of use of the different technologies is beyond the scope of this paper.

Types of Nutrition Information Capture not Considered

There are other fundamental areas in relation to capture of nutrition information such as the capture via lab testing of the various nutrient components and their proportions. However, as this is not directly an aspect of capturing the dietary or nutritional intake of individuals this will not be dealt with in detail in this paper. In addition this area of data capture is not one where informatics, mobile, computer vision or sensor techniques are the foremost candidate technologies.

Fundamentally nutritional composition is determined by food testing with data supplied by analytic efforts by government and industry, scientific literature, and from food label information (Haytowitz and Holden, 2012). Challenges for this include the variability of components; for a given food type, there is variation between different varieties e.g between different varieties of the same vegetable type, and there is variability between different brands and even different instances of a food. Partly explaining this complexity, nutrition components can arise naturally, be generated through biological stress, be added (fortification), result from manufacturing, be contaminants or be added through feeding or fertilization (Haytowitz and Holden, 2012). In addition new foods are continually becoming available and the nutrient composition of existing foods changing. Nevertheless the development of increasingly comprehensive nutrition component databases is occurring in many countries and by numerous organizations. Significant existing public initiatives include USDA SR24 in the US, EuroFIR in Europe and NUTTAB in Australia and private initiatives include the Gladson Nutrition Database. USDA SR24 for example includes over 7900 food items (Haytowitz and Holden, 2012).

IDENTIFYING FOOD TYPE, WHERE NOT A PACKAGED FOOD

Where a food is not a packaged food, substantial obstacles exist to the capture of its type and also of portion size. This includes the complex case of a home cooked meal or of a composite plate of foods, free-weight foods or foods consumed while dining out (Ng, 2012). This is the area where novel informatics, sensor, mobile, computer vision and image processing-based approaches have received most attention, are most applicable, and significant research opportunities exist.

Substantial research work has been directed to image capture and analysis of meals or plates of food to assist in identifying types and amounts of foods eaten (Sun et al., 2008; Saeki and Takeda, 2005; Martin et al., 2009a; Martin et al., 2009b; Kitamura et al., 2010; Kong and Tan, 2012; Bosch et al., 2011; Zhu et al., 2008; Zhu et al., 2010; Villabos et al., 2011, Shang et al., 2011). Lassen et al. (2010) suggest promising results in the setting of an evening meal: "the digital estimation method is a valid measure of evening meal intake with regard to macronutrient distribution, energy density, and energy-adjusted food categories compared against weighed records as the reference method". It should be noted that the digital methodology here, involved users capturing an image of their food before and after consumption, and while this is relatively convenient for research trial or dietary trial use, consideration of approaches that are even more automated for general population and ongoing usage require further investigation.

Other sensing and informatics-based research ways to make data collection more automated or requiring less manual action have been explored, with proposals such as acoustic sensing for determining the type of an individual's activity including differentiation between types of food eaten (Yatani and Truong, 2012) and acoustical dietary monitoring (Amft and Troster, 2009; Passler and Fischer, 2011; Lopez-Meyer et al., 2012; Patterson et al., 2005), specialized nutritional sensing devices (Lester et al., 2010), a multi-sensor lanyard worn around the neck (Sun et al., 2010), or smart kitchen equipment (Chang et al., 2006; Chi et al., 2008). A number of these approaches have the advantage of not requiring an individual to take such a manual action as photographing their food, but also require novel and as-yet not mass-deployed hardware. This

 ACCEPTED MANUSCRIPT

represents a challenge to the mass utilization of such novel hardware systems. Also a factor for such systems will be user perception and acceptance of utilizing sensors (Steele et al., 2009) which will have to possess sufficient usability characteristics – the acoustic and other sensors are in a number of cases wearable sensors and the smart kitchen equipment could be considered ambient sensors (Steele et al., 2009). It should however be noted, the manufacture and commercial availability of standalone wearable sensors for other consumer health-related purposes is showing rapid development at this time and there is a likely trend of the increasing integration of such emerging sensors into smartphones (Swan, 2012).

In other cases solutions are based on already widely available technology. The application of current smartphones or PDAs and the utilization of their digital camera image capture capabilities have been explored in a number of research studies (Six et al., 2010; Paldanius et al., 2011; Zhu et al., 2008; Zhu et al., 2010).

Another emerging area of research in relation to non-packaged foods that warrants further attention is the combining of computer vision or sensor approaches with informatics approaches. This would involve cross-referencing of digital image or sensor-gathered information about a food with information available in relation to what an individual or household has purchased from a grocery or other retail outlet. The combining of self-reported dietary intake information with information about what has been purchased as indicated by shopping receipts has previously been considered (Greenwood et al., 2006) with some success. However yet to be fully explored is the intelligent integration of such information input from disparate sources which can potentially improve the accuracy of the computer vision or other sensor approaches for capturing food types eaten and also portion sizes.

 ACCEPTED MANUSCRIPT

For example, when a cooked meal is captured via image, an image processing task would be to identify what type of meal this is and the food items present and their sizes. Via utilizing information about the set of foods recently purchased, information on which foods or portions are already known to have been eaten, or even recipe information, the search space of all possible meals or ingredients can be substantially constrained, to assist the computer vision-based identification of the meal and its components.

It is also acknowledged that while the goal is for such approaches to be increasingly completely automated, the use of some manual input to achieve semi-automated information capture will still provide useful benefits. Mobile app-based food journals may add to the manual input capability (Jarvinen et al, 2008; Lieffers et al., 2012; Silva et al., 2011). With the recent mass adoption of smartphones, the widespread use of mobile apps and specifically mobile health apps is a significant development for user capture of their activities and also for enabling a competitive software innovation environment in this area. The extent of the impact upon dietary information capture, of the current and future adoption of the myriad food diary and food journal mobile apps, is yet to be fully understood and researched.

DETERMINING FOOD PORTION SIZE

Capture of portion size is considered the most challenging aspect of automatically capturing dietary intake information (Martin et al., 2009b).

Particular research challenges in relation to computer vision approaches are how to carry out image recognition to determine the different foods on a plate, delimit such different foods, combine multiple images to estimate food volume and estimate amounts on a plate before and after eating (Zhu et al., 2008; Sun et al., 2008; Saeki and Takeda, 2005; Martin et al., 2009b). A sub-problem in relation to this is estimating the scale of the food image and existing approaches have included capturing with the image an image of the individual's thumb to provide a known reference for determining portion size (Villabos et al., 2011). A related and central aspect of automatically determining portion size is automatically calculating food volume or developing 3D models of the food items on a plate (Chae, et al., 2011; Yue et al., 2010).

Other approaches combine an automated aspect with manual appraisal. For example, the Remote Food Photography Method (Martin et al., 2009b) involves an individual capturing an image of their plate of food both before and after eating to allow estimation of the portion sizes actually consumed and nutritional intake. In this method, the approach involves having these images sent remotely to be analyzed by food or nutrition experts. This however raises an issue of cost for this approach and improved automated image processing-based approaches are one approach that could lower these costs by carrying out more aspects of this automatically.

In response to the cost of manual appraisal, there are also emerging novel image-based approaches such as PlateMate (Noronha et al., 2011) where the estimation of nutritional intake is derived from a 'crowd sourcing' model. That is, the expert work is sourced from a crowd of qualified individuals, thereby aiming to decrease the labour cost.

Other automated portion size capture challenges include: in the case of fluids, visually calculating how much the level of a translucent container of fluid has decreased; or in the case of a packaged food estimating what proportion of a package of food has been consumed.

There are also substantial research opportunities for further sensing approaches such as intelligent tables (Chang et al., 2006) which may be able to weigh plates of food as they are to be consumed and after eating.

CAPTURING FOOD IDENTIFIER/ UNIVERSAL PRODUCT CODE

Packaged foods can be most amenable to the task of capturing food nutritional content and intake, as when a packaged food with a food identifier is consumed in whole or in a serving size proportion, the detailed nutrition intake information is typically immediately known through nutrient component databases or a nutrition fact panel, and all that needs to be inputted is the packaged food's identifier.

Fundamentally the availability of a food identifier and its capture visually or via other means is particularly powerful in relation to capturing dietary and nutritional intake information. The most common food identifiers and widely utilized at this time are Universal Product Codes (UPCs) which are typically available as bar codes on food packages. These bar codes currently are inputted via scanning by those in the food production and retail industry and also increasingly can be scanned by consumers using smartphone-based digital cameras, other cameras, scanners or wireless communication mechanisms. In addition another significant international identifier standard is the Global Trade Item Number (GTIN) from which the GS1 system was originated (http://www.gs1.org/). There are also other identifier systems relevant to foods other than barcodes and these include QR Codes and proprietary labels.

The benefits of a UPC are that it unambiguously identifies the food, is typically accompanied by a nutrition fact panel, subsequently enables unique lookup of other more detailed nutrient databases (that are indexed by UPC) for the detailed nutrition information corresponding to that UPC (Slining, 2012) and it also can powerfully help to assist the addressing of the question of portion size eaten. That is, where the whole amount of a packaged food is eaten, or the recommended serving, or an estimated fraction of the packaged food, there then is a simple step to calculate dietary and nutritional intake consumed from that food.

The automation challenges in capturing this information are substantially less than for some of the other listed types of nutrition intake data capture covered previously in this paper. As such, approaches based on input of food identifiers are also further amenable to utilization on a general population scale. Scanning of barcodes is an existing capability including for home use (Stevens et al, 2010). In the case of other food identifier types the input of the identifier via computer vision approaches is again already an existing capability such as in the case of QR Codes or more basically via Optical Character Recognition (OCR). The main challenge will still be estimating portion size consumed from a packaged food, and this has been described in more detail in a previous section.

There is scope for other sensing capabilities to play an important role in improved food identifier capture. Sensing capabilities in relation to food identifiers include using non-image based approaches to capturing the identifier such as RFID-based approaches (Philipose et al., 2004). Such approaches may have benefits of usability over image-based approaches, as they do not require the positioning of a smartphone, camera or scanner to capture an image of the barcode. Such approaches as RFID would just require proximity. On the other hand, the practicality of the use of RFID on all food items or the mass deployment of RFID scanners represents real impediments to this approach.

Capturing the food identifier or scanning could be done either while shopping, at time of check-out or at home. Wireless approaches may offer some promise. The possibility for the itemized purchase receipt information to be transmitted to an individual's smartphone by wireless or via other means at point of purchase or via other means is a feasible and realistic functionality development. Another possibility is detailed purchase records, including for example more detailed credit card data available to the consumer. These could allow the automatic inputting of food items purchased and hence allow the determination of aggregate nutritional content.

Finally it should be noted that the capture of UPC information may not remain a sensor or computer vision task. Given that this information is available in a structured format it is likely to become a fundamental integration capability of retail food systems or nutrition informatics systems.

It should be noted however that the widespread adoption of such retail nutrition information systems is at least a number of years away or possibly more. In the meantime, sensor or image scanning approaches to retrieve food identifier information may remain the most practical, and in some cases, only currently available mechanism.

A further implication of the above possibilities for improved integration of retail information systems is that image-based inputting of food identifiers may be most important in determining the nutritional content of food items before purchase. Check-out or credit card based approaches may provide mechanisms in future to capture your full nutritional intake that requires no manual effort by the consumer, including no manual scanning or image capture effort. The important role of scanning prior to purchase occurs when a consumer is choosing between possible

 ACCEPTED MANUSCRIPT

products to purchase and may wish to retrieve more detailed nutrition information to assist in that decision making.

Capture of food purchase and consumption data has traditionally been implemented for market research, statistics gathering or research purposes. Services such as Scantrack and Homescan by Nielsen and What We Eat in America (WWEIA-NHANES) by the USDA and US-CDC provide survey-based data (Slining, 2012). However there are significant limitations to these traditional survey-based approaches to collecting end-user consumption information, in particular they are limited in terms of numbers participating, limited to capturing household purchases and most importantly do not involve capture of an individual's actual food and nutrition intake nor provide such information for the individual to use. These tools have not had the goal of supporting consumer self-knowledge of dietary and nutritional intake and so have not provided any form of nutritional guidance service for individuals.

Efforts to extend the usage of food UPCs to be more universal, and also increasingly in eat-out menus (Nestle, 2010), will advance the capture of food consumption information and this represents a research area where policy, adoption and stakeholder issues will play a part. Policy change to introduce and require universal provision of food identifiers represents a complementary and potentially more powerful pathway to achieving more capable and sophisticated automated dietary intake information capture and the ability of all in the population to have a complete knowledge of their detailed nutritional intake.

While UPC identifiers have been adopted to address the tracking needs of retailers, there would be benefits from a more systematic coding of all food types, varieties and items that is not driven by just retailing considerations. For drugs there are National Drug Code (NDC) codes (FDA, 2012), but there are no such standard codes for foods. Potentially relevant existing codifications that may provide a starting point towards this include the Unique Ingredient Identifier codes (UNII) (NIH, 2012) and the Langual food description thesaurus (http://www.langual.org/). International efforts to produce further standardization in this regard would be beneficial.

NUTRITION FACTS PANELS

Another relevant type of data available for capturing dietary intake information is the nutrition facts panel. Capture of such information is also much more trivial for computer vision approaches than determining food type for unpackaged foods or determining portion size. Such labels are required in many circumstances and countries by law (Nestle, 2010). These labels typically provide such information as macronutrient content amounts, proportions of carbohydrates, protein and fats and often proportions of a number of other nutritional components.

In relation to the application of informatics-based approaches to nutrition facts panels, a challenge relates to how these can be displayed in a customized, more detailed or less detailed way to individuals. Interface issues will be an important part of how these labels can be understood when presented in a customized or personalized way to individuals.

As with the case of food identifiers and UPCs, there may emerge alternative approaches which utilize more integrated nutrition information systems. For example the nutrition facts panel information can be retrieved from a remote data base of such nutrition panel information, once

 ACCEPTED MANUSCRIPT

the UPC has been scanned or otherwise inputted. This would also decrease the role of computer vision or sensor-based approaches for this particular type of dietary intake-related data capture.

However given that such integration is currently not generally available at this time, inputting of the nutrition facts panel via computer vision or OCR-based approaches remains a currently available approach. Utilizing a knowledge of the standard format and structure of nutrition facts panels can help to cross-validate and improve the accuracy of such computer vision-based approaches.

RECIPE INFORMATION

Recipes captured in digital form also represent a powerful adjunct information source for assisting in automating the capture of nutritional intake information, where individuals are eating meals that have been cooked or prepared. It should be noted that there are various Web-based and software-based approaches that have supported the automated capture of recipe information (Li et al., 2006).

While not extensively developed there is the potential for image processing and sensor systems that capture food amounts used during cooking or food preparation. This is an avenue of research that can draw upon smart kitchen based technologies (Chi et al., 2008).

Recipe information will also be an important input for calculating the nutritional content of dine-out meals. Recipe information in these circumstances could assist a dine-out business to calculate the nutritional content of their own foods, with the nutritional content then made available to those who dine at their establishment.

Recipes, with corresponding ingredient quantities are able to assist in the determination of the aggregate nutritional content of a given prepared food or meal. Of course it is acknowledged that despite the uptake of electronic recipe systems, in many cases there will not be recipe information available or available in an electronic format for a given home-prepared food or meal.

But where recipe information if available, recipe information can also be cross-correlated with information about recent individual or household food purchases to further delineate the content of a given prepared meal. This cross-correlation can be used to improve the accuracy once again of the image processing-based approaches to identifying foods and meal components. This area has also been discussed in the preceding section.

AGGREGATE APPROACHES

Another possibility is that an individual's itemized purchase record can be automatically fed into their record of dietary intake (Mankoff et al., 2002; Brinkerhoff et al., 2011) and via this means allow for individuals to calculate their aggregate nutritional intake.

As described above, the capture of exact food types and exact portion sizes, in some cases, will continue to be challenging obstacles, even though mobile computing, sensor and computer vision-based approaches can be applied to further these capabilities. Another approach that can partly obviate these challenges is capturing aggregate food purchase and intake information. For example an individual's total grocery load can be inputted at time of purchase via UPC information (Brinkerhoff et al, 2011) or later (Stevens et al., 2010) and this can provide information about what that individual eats over the next period of time including detailed aggregate nutrition information included through integration with nutrition component databases (Brinkerhoff et al., 2011; Slining, 2012). In many cases this will be sufficient to know an individual's nutrition or at least food use patterns (Weinstein et al., 2006), even if the specific proportions eaten in any given meal might not be known. Calculation of the period over which this food was eaten can be facilitated, for example, by recording when each food item is re-purchased, using the heuristic that the food is typically re-bought when that previously bought item is finished or near-finished.

A computer vision approach to this can involve scanning of shopping receipts (Mankoff et al., 2002). While this is still manually tedious it does provide numerous benefits in calculating nutritional intake. Improved systems can however improve in relation to the manual effort required including via wireless information transfer, for example at the point of purchase or via integration with itemized digital purchase records. The use of smartphone cameras alternatively offers a pervasive and readily available means for scanning.

Such an aggregate approach will be most effective in one-person houses. In houses of more than one member combining this aggregate approach with some level of image capture or sensor-based approach can help in determining which house members eat which food items and portions. However, where within a house prepared meals are shared, the total aggregate nutritional intake can be estimated from the aggregate purchases. Proportions eaten from a common prepared meal could then be estimated either manually or via any of the approaches and technologies described previously in relation to determining portion size.

ESTIMATION APPROACHES

As indicated earlier, capture of exact food type including all ingredient information and portion sizes for all meals will be technically challenging and even where there is significant automation is likely to involve some user inconvenience.

Estimation techniques that could be utilized can include: sampling - capturing intake for some meals and using these to estimate overall intake; repeated consumption - many food item types are eaten repeatedly by an individual so nutrition information for an item may need to be captured just once and then can be re-used; consumption patterns - individuals and members of multi-member households follow regular consumption patterns and this can be exploited to save information capture effort; and portion sizes eaten will also display regularity.

As more data about an individual's dietary intake is captured this can be used to assist in future estimation or extrapolation to reduce the need for future data input and particularly manual data input.

SUMMARY AND PROMISING DIRECTIONS

We have described a wide range of technologies that are currently available for the automated capture of dietary intake and nutritional intake information. Substantial research work has examined automated image analysis of food items or plates of food eaten and significant progress has been made in this area. However image analysis approaches still first require the capturing of a digital image of the food to be eaten, which presents challenges for utilization by the general public in automating dietary intake capture. Other wearable and ambient sensor-based approaches have also been discussed that may remove the need for any manual action by an individual such as image capture to take place, but these sensors also introduce in some cases

 ACCEPTED MANUSCRIPT

the issues of the requirement of novel, non-commodity hardware, the level of sensing capabilities and user acceptance and adoption of such sensor systems. Smartphones are however incorporating more built-in sensors and sensors that wirelessly connect to smartphones are increasingly becoming available as affordable consumer electronics products.

At the same time, there is the increasing capability to automatically capture food purchase information and grocery receipt information in digital form in a way that requires little or no manual intervention and via this information to calculate the full aggregate nutritional content of such food purchases. The capture of such information is being continually facilitated by the greater digitization of point-of-sale food retail systems, the wider usage of barcodes and food identifiers, the ability to calculate nutritional content via the increasing availability and sophistication of nutrition component databanks indexed by food item, and the increasing requirement for and prevalence of nutrition facts panels.

A promising direction for further development of automation in the capture of dietary intake information is the further integration of the information from all of the numerous disparate information sources mentioned above. For example, food purchase information can assist in the disambiguation of visual or sensor-based data in relation to what an individual actually consumes. Or image or sensor data indicating information about what an individual actually consumes can be used to segment aggregate food and nutrition totals determined from purchase information, for example for a household, into portions and proportions eaten by different individuals in that house.

Further, automated capture of dietary intake information can be further facilitated by the prioritization of the following developments: greater adoption of food identifiers and barcodes including in digital form, further data standardization in relation to the digital exchange of dietary and nutrition information to support integration between and amongst consumer and retailer information systems, greater access to and integration with nutrition component databases, more sophisticated nutrition component databases, more sophisticated, nutrition facts panels including in an electronic form and the development of image capture and sensor and hardware capabilities that require less manual effort.

CONCLUSION

In this paper the state of the art in automatic dietary and nutrition information capture technologies has been overviewed and categorized to describe the applications of informatics, mobile computing, computer vision and sensor-based approaches to the capture of the following types of dietary intake-related information: food type, portion size, food identifiers, recipe information, nutrition facts panels and aggregate itemized purchase information. How the capture of this information contributes to overall automated dietary intake information capture is discussed throughout.

In summary, systems to automatically make available total and itemized food purchase information from the point-of-sale, and integrated with nutrition component databases have the potential to automate the capture and calculation of overall aggregate dietary intake at a household or in some cases individual level, in a way ideally minimizing the requirement for manual actions by individuals. In conjunction, smartphones, or future variant devices, via their built-in cameras and sensors, and potentially via the numerous external sensor capabilities being

 ACCEPTED MANUSCRIPT

researched, developed and in some cases already marketed, are most promising to provide the essential information to assist resolving this aggregate information into dietary intake information at the individual and individual sittings levels of granularity. Of those smartphone and sensor systems, those that require less manual action by the user will tend to be more readily adopted. At the same time, the increasingly ubiquitous use of food identifiers and an infrastructure to integrate corresponding and more extensive and sophisticated electronically-stored nutrition-related data with these identifiers, will work to advance and make easier all aspects of automated dietary and nutrition intake capture. Ultimately a combination of the technologies discussed will be needed, not just one single technology.

In addition related outstanding research problems have been identified. An outstanding research area receiving substantial attention is improving image analysis of digital images of food, either carried out automatically or manually. At the same time there is also substantial and ongoing research into other sensor technology approaches such as audio-based, lanyards, smart tabletop and smart kitchen technologies to also automatically capture dietary intake that may obviate the need for the manual capture of images of food. There is also research into the capturing of shopping receipt information via scanning or other means, and research into nutrition component databases and emerging research into integration of such information sources.

The most promising directions include: further attention on software systems, techniques and architectures involving the integrating together of itemized purchase information with the other disparate types of food and nutrition intake related information; sensor-based systems not requiring manual action by the user; greater development of international food and nutrition data standards and identifiers; research into adoption and diffusion of such standards; mobile apps that further engage the consumer in their dietary choices; and improved computer vision techniques to better analyze food-related images.

REFERENCES

Amft, O. and Troster, G. (2009). On-body sensing solutions for automatic dietary monitoring. IEEE Pervasive Computing 8, 2 : 62-70.

Barreiro-Hurle, J., Gracia, A., and de-Magistris, T. (2010). Does nutrition information on food products lead to healthier food choices? Food Policy 35(3): 221-29.

Blumenthal, K. and Volpp, K.G. (2010). Enhancing the effectiveness of food labelling in restaurants. Journal of the American Medical Association, 303(6): 553-554.

Bosch, M., Zhu, F., Khanna, N., Boushey, C. J., & Delp, E. J. (2011). Combining global and local features for food identification in dietary assessment. In Image Processing (ICIP), 2011 18th IEEE International Conference on (pp. 1789-1792). IEEE.

Brinkerhoff, K., Brewster, P., Clark, E., Jordan, K., Cummins, M. and Hurdle, J. (2011). Linking supermarket sales data to nutritional information: an informatics feasibility study. AMIA Annu Symp Proc. 2011: 598–606.

Burke, L.E., Wang, J., Sevick, M.A. (2011). Self-monitoring in weight loss: a systematic review of the literature. Journal of the American Dietetic Association, 111(19): 92–102.

Chae, J., Woo, I., Kim, S., Maciejewski, R., Zhu, F., Delp, E. J., ... & Ebert, D. S. (2011). Volume estimation using food specific shape templates in mobile image-based dietary assessment. In Proceedings of SPIE (Vol. 7873, p. 78730K). NIH Public Access.

Chang, K.-h., Liu, S.-y., Chu, H.-h., Hsu, J. Y.-j., Chen, C., Lin, T.-y., Chen, C.-y., and Huang, P. (2006). The diet-aware dining table: observing dietary behaviors over a tabletop surface.

Proceedings of the 4th international conference on Pervasive Computing, ser. PERVASIVE'06. (pp. 366–382). Berlin, Heidelberg: Springer-Verlag.

Chatterjee, S. and Price, A. (2009). Healthy living with persuasive technologies: framework, issues, and challenges. J Am Med Inform Assoc, 16(2):171–8.

Chi, P.-Y., Chen, J.-H., Chu, H.-H. and Lo, J.-L. (2008). Enabling calorie-aware cooking in a smart kitchen, in Persuasive Technology, ser. Lecture Notes in Computer Science, H. Oinas-Kukkonen, P. Hasle, M. Harjumaa, K. Segersthl, and P. hrstrm, Eds. Springer Berlin / Heidelberg, 2008, vol. 5033: 116–127.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K. and Dixon, D. (2011). Gamification: using game-design elements in non-gaming contexts. Proceedings of the CHI extended abstracts:.2425 – 2428.

Donnelly, K., van der Roest, J., Hoskuldsson, S. and Olsen, P. (2011). Food industry information exchange and the role of meta–data and data lists. International Journal of Metadata, Semantics and Ontologies, 6(2): 146-153.

Dutton, W.H. and Eynon, R. (2009). Networked individuals and institutions: a cross-sector comparative perspective on patterns and strategies in government and research, The Information Society, 25 (3): 1-11.

Finch, C.E. (2007). The Biology of Human Longevity: Inflammation, Nutrition, and Aging in the Evolution of Lifespans (Academic, San Diego).

Fogg, BJ. (2003). Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann, San Francisco.

Food and Drug Administration, US. (2012). National Drug Code Directory. Retrieved 16 June 2012, from http://www.fda.gov/Drugs/InformationOnDrugs/ucm142438.htm.

Greenwood, D. C., Ransley, J. K., Gilthorpe, M. S., & Cade, J. E. (2006). Use of itemized till receipts to adjust for correlated dietary measurement error. American journal of epidemiology, 164(10): 1012-1018.

Hawkes, C. (2012). Food policies for healthy populations and healthy economies. BMJ 344:e2801.

Haytowitz, D. and Holden, J. (2012). USDA food composition and nutrient databases. Proceedings of the 36th National Nutrient Databank Conference. Retrieved 16 June 2012, from http://www.nutrientdataconf.org/PastConf/NDBC36/W-1_Haytowitz-Holden_NDDC2012.pdf.

Illner, A. K., Freisling, H., Boeing, H., Huybrechts, I., Crispim, S. P., & Slimani, N. (2012). Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. International Journal of Epidemiology, 41(4): 1187-1203.

Jarvinen, P., Jarvinen, T. H., Lahteenmaki, L., & Sodergard, C. (2008). HyperFit: hybrid media in personal nutrition and exercise management. In Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. Second International Conference on (pp. 222-226). IEEE.

Jordan, E. T., Ray, E. M., Johnson, P. and Evans, W. D. (2011). Text4Baby. Nursing for Women's Health, 15: 206–212. doi: 10.1111/j.1751-486X.2011.01635.x

Karvetti, R. L. and Knuts, L.R. (1985). Validity of the 24-hour dietary recall. J Am Diet Assoc 185 : 1437-42.

 ACCEPTED MANUSCRIPT

Kitamura, K., de Silva, C., Yamasaki, T. and Aizawa, K. (2010). Image processing based approach to food balance analysis for personal food logging. In IEEE International Conference on Multimedia and Expo (ICME). IEEE : 625–630.

Klasnja, P., Consolvo, S., McDonalid, D., Landay, J. and Pratt, W. (2009). Using mobile & personal sensing technologies to support health behavior change in everyday life: lessons learned. AMIA Annu Symp Proc. 2009: 338–342.

Kong, F., & Tan, J. (2012). DietCam: automatic dietary assessment with mobile camera phones. Pervasive and Mobile Computing, 8(1): 147-163.

Lassen, A. D., Poulsen, S., Ernst, L., Andersen, K. K., Biltoft-Jensen, A., & Tetens, I. (2010). Evaluation of a digital method to assess evening meal intake in a free-living adult population. Food & Nutrition Research, 54.

Lester, J., Tan, D., Patel, S. and Brush, A. (2010). Automatic classification of daily fluid intake, in Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on,: 1 –8.

Li, Y., Meng, X., Wang, L. and Li, Q. (2006). RecipeCrawler: collecting recipe data from WWW incrementally. In Proc. of the 7th International Conference on Web-Age Information Management (WAIM), Hong Kong, China: 263-274.

Lieffers, J. R., & Hanning, R. M. (2012). Dietary assessment and self-monitoring with nutrition applications for mobile devices. Canadian Journal of Dietetic Practice and Research, 73(3), 253-260.

 ACCEPTED MANUSCRIPT

Lindqvist, A-K., Kostenius, C., and Gard, G. (2012). Peers, parents and phones – Swedish adolescents and health promotion. Int J Qual Stud Health Well-being. 2012; 7: 10.3402/qhw.v7i0.17726.

Lopez-Meyer, P., Schuckers, S., Makeyev, O., Fontana, J.M. and Sazonov, E. (2012). Automatic identification of the number of food items in a meal using clustering techniques based on the monitoring of swallowing and chewing. Biomedical Signal Processing and Control, 7(5): 474 – 480.

Mankoff, J., Hsieh, G., Hung, H.C., Lee, S. and Nitao, E. (2002). Using low-cost sensing to support nutritional awareness. In: Goos G, Hartmanis J, van Leeuwen J, editors. Ubicomp 2002: Proceedings of the 4th International Conference on Ubiquitous Computing, vol. 2498 of Lecture Notes in Computer Science, Springer Berlin, Heidelberg: 371—6.

Martin, C., Kaya, S. and Gunturk, B. (2009a). Quantification of food intake using food image analysis, Conf Proc IEEE Eng Med Biol Soc., 6869–6872.

Martin, C., Han, H., Coulon, S., Allen, H., Champagne, C. and Anton, S. (2009b). A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. British Journal of Nutrition, 101, 03: 446–456.

Maruthur, N.M., Wang, N.Y., and Appel, L.J. (2009). Lifestyle interventions reduce coronary heart disease risk: results from the PREMIER trial. Circulation, 119(15): 2026-31.

Mutchler, L., Shim, J., and Ormond, D. (2011). Exploratory study on users' behavior: smartphone usage. Proceedings of AMCIS 2011. (pp. 10). AIS Electronic Library (AISeL).

Nestle, M. (2010). Health care reform in action – calorie labeling goes national. New England Journal of Medicine. 362**:** 2343-2345.

 ACCEPTED MANUSCRIPT

National Institutes of Health. (2012). Unique Ingredient Identifier (UNII). Retrieved 16 June 2012, from http://fdasis.nlm.nih.gov/srs/srs.jsp.

Ng, S.W. (2012). Addressing gaps in market level databases, Proceedings of the 36th National Nutrient Databank Conference, 2012. Retrieved 16 June 2012, from http://www.nutrientdataconf.org/PastConf/NDBC36/5-3_Ng_NNDC2012.pdf

Ngo, J., Engelen, A., Molag, M., Roesle, J., García-Segovia, P., & Serra-Majem, L. (2009). A review of the use of information and communication technologies for dietary assessment. Br J Nutr, 101(Suppl 2), S102-12.

Noronha, J., Hysen, E., Zhang, H. and Gajos, K. Z. (2011). Platemate: crowdsourcing nutritional analysis from food photographs. In Proc. UIST 2011, ACM: 1-12.

Paldanius, M., Lehtiö, A., Karukka, M., & Huuskonen, P. (2011). Personal chart: health data logging made easy with a mobile device. Human-Computer Interaction–INTERACT 2011, 592-595.

Passler, S. and Fischer, W.-J. (2011). Acoustical method for objective food intake monitoring using a wearable sensor system, in Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on, may 2011: 266 –269.

Patterson, D.J. et al. (2005). Fine-grained activity recognition by aggregating abstract object usage, Proc. 9th IEEE Int'l Symp. Wearable Computers: 44–51.

Philipose, M., Fishkin, K.P., Perkowitz, M., Patterson, D., Fox, D., Kautz, H. and Hahnel, D. (2004). Inferring activities from interactions with objects. IEEE Pervasive Computing 3, 4: 50-57.

 ACCEPTED MANUSCRIPT

Saeki, Y. and Takeda, F. (2005). Proposal of food intake measuring system in medical use and its discussion of practical capability, LNCS, Volume 3683/2005 : 181-183.

Shang, J., Sundara-Rajan, K., Lindsey, L., Mamishev, A., Johnson, E., Teredesai, A., & Kristal, A. (2011). A pervasive dietary data recording system. In Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on (pp. 307-309). IEEE.

Shay, L.E., Seibert, D., Watts, D., Sbrocco, T., and Pagliara, C. (2009). Adherence and weight loss outcomes associated with food-exercise diary preference in a military weight management program. Eating Behaviors, 10(4), 220–227.

Shils, M.E., ed, Shike, M., ed, Ross, A.C., ed, Caballero, B., ed, Cousins, R.J., ed. (2006). Modern Nutrition in Health and Disease. 10th ed. Philadelphia, PA: Lippincott Williams and Wilkins, 982.

Silva, B. M., Lopes, I. M., Rodrigues, J. J., & Ray, P. (2011). SapoFitness: A mobile health application for dietary evaluation. In e-Health Networking Applications and Services (Healthcom), 2011 13th IEEE International Conference on (pp. 375-380). IEEE.

Six, B., Schap, T., Zhu, F., Mariappan, A., Bosch, M., Delp, E., Ebert, D., Kerr, D. and Boushey, C. (2010). Evidence-based development of a mobile telephone food record. J Am Diet Assoc. 2010 January; 110(1): 74–79.

Slining, M. (2012). Linking data sources. Proceedings of the 36th National Nutrient Databank Conference. Retrieved 16 June 2012, from http://www.nutrientdataconf.org/PastConf/NDBC36/5-6_Slining_NNDC2012.pdf.

Steele, R. and Clarke, A. (2012). A real-time, composite healthy building measurement architecture drawing upon occupant smartphone-collected data. Proceedings of the 10th International Healthy Buildings Conference.

Steele, R., Lo, A., Secombe, C., Wong, Y.K. (2009). Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare, International Journal of Medical Informatics 78 (12) : 788–801.

Steele, R., Min, K. and Lo, A. (2012). Personal health record architectures: technology infrastructure implications and dependencies. Journal of the American Society for Information Science and Technology, 63(6), 1079-1091.

Stevens, J., Bryant, M., Wang, L., Borja, J., & Bentley, M. E. (2010). Exhaustive measurement of food items in the home using a universal product code scanner. Public Health Nutrition, 14(2), 314.

Strauss, J. and Thomas, D. (1998). Health, nutrition and economic development. Journal of Economic Literature, 36(6), 766-817.

Sturm, R. (2002). The effects of obesity, smoking and drinking on medical problems and costs. Health Affairs, 21(2), 245- 253.

Sun, M., Fernstrom, J. D., Jia, W., Hackworth, S. A., Yao, N., Li, Y., & Sclabassi, R. J. (2010). A wearable electronic system for objective dietary assessment. Journal of the American Dietetic Association, 110(1), 45.

Sun, M., Liu, Q., Schmidt, K., Yang, J., Yao, N., Fernstrom, J., Fernstrom, M., DeLany, J. and Sclabassi, R. (2008). Determination of food portion size by image processing, 30th Annual International IEEE EMBS Conference: 871 - 874.

Swan, M. (2012). Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0, Journal of Sensor and Actuator Networks 1 (3): 217-253.

Villalobos, G., Almaghrabi, R., Hariri, B. and Shirmohammadi, S. (2011). A personal assistive system for nutrient intake monitoring, in Proceedings of the 2011 international ACM workshop on Ubiquitous meta user interfaces, ser. Ubi-MUI ’11. New York, NY, USA: 17–22.

Weinstein, J. L., Phillips, V., MacLeod, E., Arsenault, M., & Ferris, A. M. (2006). A universal product code scanner is a feasible method of measuring household food inventory and food use patterns in low-income families. Journal of the American Dietetic Association, 106(3): 443-445.

WHO. (2003). Diet, Nutrition and the Prevention of Chronic Diseases. Report of a Joint WHO/FAO Expert Consultation. WHO Technical Report Series 916. Geneva: World Health Organisation.

Yatani, K. and Truong, K. (2012). BodyScope: a wearable acoustic sensor for activity recognition. Proceedings of UbiComp, Sept. 2012.

Yue, Y., Jia, W., Fernstrom, J. D., Sclabassi, R. J., Fernstrom, M. H., Yao, N., & Sun, M. (2010). Food volume estimation using a circular reference in image-based dietary studies. In Bioengineering Conference, Proceedings of the 2010 IEEE 36th Annual Northeast (pp. 1-2). IEEE.

Zhu, F., Bosch, M., Woo, I., Kim, S., Boushey, C. J., Ebert, D. S., & Delp, E. J. (2010). The use of mobile devices in aiding dietary assessment and evaluation. Selected Topics in Signal Processing, IEEE Journal of, 4(4), 756-766.

Zhu, F., Mariappan, A., Kerr, D., Boushey, C., Lutes, K., Ebert, D. and Delp, E. (2008). Technology-assisted dietary assessment, Proceedings of the IS&T/SPIE Conference on Computational Imaging VI, vol. 6814.