Human-centered Computing

Nicu Sebe

1 Introduction

Computing is at one of its most exciting moments in history, playing an essential role in supporting many important human activities. The explosion in the availability of information in various media forms and through multiple sensors and devices means, on one hand, that the amount of data we can collect will continue to increase dramatically, and, on the other hand, that we need to develop new paradigms to search, organize, and integrate such information to support all human activities.

Whether we talk about the pervasive, ubiquitous, mobile, grid, or even the social computing revolution, we can be sure that computing is impacting the way we interact with each other, the way we design and build our homes and cities, the way we learn, the way we communicate, play, or work [52][61]. Simply put, computing technologies are increasingly affecting and transforming almost every aspect of our daily lives.

Unfortunately, the changes are not always positive and much of the technology we use is clunky, unfriendly, unnatural, culturally biased, and difficult to use. As a result, several aspects of daily life are becoming increasingly complex and demanding.

We have access to huge amounts of information, much of which is irrelevant to our own local socio-cultural context and needs or is inaccessible because it is not available in our native language, we cannot fully utilize the existing tools to find it, or such tools are inadequate or nonexistent. Thanks to computing technologies, our options for communicating with others have increased, but that does not necessarily mean that our communications have become more efficient. Furthermore, our interactions with computers remain far from ideal and too often only literate,

Nicu Sebe
University of Amsterdam, The Netherlands, e-mail: nicu@science.uva.nl
& University of Trento, Italy, e-mail: niculae.sebe@unitn.it

H. Nakashima et al. (eds.), *Handbook of Ambient Intelligence and Smart Environments*,
DOI 10.1007/978-0-387-93808-0_13,

educated individuals who invest significant amounts of time in using computers can take direct advantage of what computing technologies have to offer.

The goal of this chapter is to discuss the most important aspects of HCC covering the existing challenges (Section 2.1), the different possible definitions of HCC reflecting the wide expertise of the researchers working in this field, the scope of HCC, and the connection with related research areas (Sections 2.2, 2.3, and 2.4, respectively). We also present the areas of research covered by HCC and we focus on three main directions: production, analysis, and interaction (Section 3). We then present the possible applications (Section 4) and we discuss the issues of integrating the existing systems and applications into real world conditions (Section 5). Finally, we propose a research agenda for HCC (Section 6).

2 What is Human-centered Computing?

Human Centered Computing (HCC) is an emerging field that aims at bridging the existing gaps between the various disciplines involved with the design and implementation of computing systems that support human's activities. HCC aims at tightly integrating human sciences (e.g. social and cognitive) and computer science (e.g. human-computer interaction (HCI), signal processing, machine learning, and ubiquitous computing) for the design of computing systems with a human focus from the beginning to the end. This focus should consider the personal, social, and cultural contexts in which such systems are deployed [57]. Beyond being a meeting place for existing disciplines, HCC also aims at radically changing computing with new methodologies to design and build systems that support and enrich people's lives.

2.1 Gateways and Barriers

We could argue that knowledge and communications are two of the main pillars of any society. As information of all kinds increasingly forms part of the digital ecosystem, the computing technologies we develop become, paradoxically, both the gateways to all kinds of resources and the barriers to access them. Therefore, in addition to having become critical in improving lives, computing is now essential to the livelihood not only of the relatively few of us who have the necessary skills and access to resources, but potentially for everyone [8][38].

For the most part, however, the computing community designs and implements computing algorithms and technologies without fully taking into account our cognitive abilities, the ways we perceive and handle information, go about our work and life, create and maintain our social relations, or use our cultural context. That is, researchers and engineers often develop computing technologies in relative isolation.

Most current methodologies start with an idea that builds on or improves existing technologies or that solves problems in a particular technological domain, largely ignoring human issues. The obvious outcomes are more powerful, less expensive computers that are more difficult to use, in some cases even effectively slower and less accessible to most of the world's population. Given the current rate of penetration of computing technologies (directly or indirectly) in almost every imaginable human activity, it is clear that the existing computing research and development models are no longer adequate. People who adapt more quickly to technology (whether the technology is good or bad) have greater opportunities to benefit at many levels.

Developing technologies in relative human isolation does not contribute to alleviating the problems of wealth distribution, sustainability, and access to healthcare and education. Technologies that are difficult to use not only waste our time and make our lives worse, they make access to important resources more difficult, particularly for people who do not have the education or language skills of the minority who develop those technologies.

From this perspective, the current path of development of computer technologies is not only detrimental but also dangerous because it contributes to increasing the gap between the educated and uneducated, and between the rich and the poor. The problem is even greater if we consider that the difficulties in using much of the technology available today are leading to a rapidly increasing content gap - most digital information is produced in developed countries, from particular cultural perspectives, and in only a few languages. To make things worse, access to the information itself is through technology developed without considering the local socio-cultural context of the majority of the world's population.

Given the developments in recent years in decreased computing costs, increased wireless communications, and the spread of the Internet, the time is ripe for a major shift in the computing revolution. In our opinion, the goal of human-centered computing is to focus the computing revolution on addressing human abilities and needs.

2.2 Definitions

In the last few years, many definitions of HCC have emerged. In general, the term HCC is used as an umbrella to embrace a number of definitions which were intended to express a particular focus or perspective [29]. In 1997, the U.S. National Science Foundation supported a workshop on Human-Centered Systems [19], which included position papers from 51 researchers from various disciplines and application areas including electronics, psychology, medicine, and the military. Participants proposed various definitions for HCC, including the following (see [19]):

- HCC is "a philosophical-humanistic position regarding the ethics and aesthetics of the workplace";
- a HCC system is "any system that enhances human performance";

- a HCC system is “any system that plays any kind of role in mediating human interactions”;
- HCC is “a software design process that results in interfaces that are really user-friendly”;
- HCC is “a description of what makes for a good tool - the computer does all the adapting”;
- HCC is “an emerging inter-discipline requiring institutionalization and special training programs”.

Other definitions of HCC have also appeared in the literature:

- According to [20], HCC is “the science of designing computations and computational artifacts in support of human endeavors”;
- For Canny et al. [10], HCC is “a vision for computing research that integrates technical studies with the broad implications of computing in a task-directed way. HCC spans computer science and several engineering disciplines, cognitive science, economics and social sciences.”

So, what is really HCC? HCC research focuses on all aspects of human-machine integration: humans with software, humans with hardware, humans with workspaces, humans with humans, as well as aspects of machine-machine interaction (e.g., software agents) if they impact the total performance of a system intended for human use. The HCC vision inherits all of the complexity of software engineering and systems integration with the human in the loop, plus the additional complexity of understanding and modeling human-human and human-computer interaction in the context of the working environment [13].

HCC recognizes the fact that the design and the use of new information processing tools must be couched in system terms [19]. That is, the humans and the information processing devices are regarded as a coupled, co-adapting system nested within a context in which they are functional. Furthermore, the design of systems must regard the human as just one aspect of a larger and dynamic context, including the team, organization, work environment, etc. This means that the fundamental unit of analysis is not the machine, nor the human, nor the context and domain of work, but the triple including all three [19]. In this context, according to Hoffman [28], HCC can be defined as “the development, evaluation, and dissemination of technology that is intended to amplify and extend the human capabilities to:

- perceive, understand, reason, decide, and collaborate;
- conduct cognitive work;
- achieve, maintain, and exercise expertise.”

Inherently, HCC research is regarded as being interdisciplinary, as illustrated by the participation of experts from a wide range of fields, including computer, cognitive, and social sciences in defining HCC and its scope.

Based on these observations we adopt the following definition: Human-Centered Computing, more than being a field of study, is a set of methodologies that apply to any field that uses computers, in any form, in applications in which humans directly interact with devices or systems that use computer technologies.

2.3 Scope of HCC

One way to identify the scope of HCC is to examine new and existing initiatives in research and education. Given the trends that indicate that computing is permeating practically all areas of human activity, HCC has received increasing attention in academia as a response to a clear need to train professionals and scholars in this domain. A number of initiatives have appeared in the past years. Examples of the growing interest in HCC are the HCC doctoral program at Georgia Tech [27], the HCC consortium at the University of California, Berkeley [26], and the Institute of Human and Machine Cognition, Florida [30], to mention just a few.

The goals are ambitious. For instance, Georgia Tech's interdisciplinary Ph.D. program, created in 2004, "aims to bridge the gap between technology and humans by integrating concepts from anthropology, cognitive sciences, HCI, learning sciences and technology, psychology and sociology with computing and computer sciences [...] with the explicit goal of developing theory and experimentation linking human concerns and computing in all areas of computing, ranging from the technical-use focus of programming languages and API designs and software engineering tools and methodologies to the impacts of computing technology on individuals, groups, organizations, and entire societies" [20]. The program emphasizes, on one hand, a deep focus on "theoretical, methodological, and conceptual issues associated with humans (cognitive, biological, socio-cultural); design; ethics; and analysis and evaluation", and, on the other hand, "design, prototyping, and implementation of systems for HCC", with a focus on building interactive system prototypes.

Conceiving computing as the "infrastructure around a human activity", the UC, Berkeley HCC consortium is "a multidisciplinary program designed to guide the future development of computing so as to maximize its value to society" [10]. The initiative added a socio-economic dimension to the domain: "the great economic [computing] advances we have seen are undermined because only a fraction of the population can fully use them. Better understanding of human cognition is certainly needed, but also of the social and economic forces that ubiquitous computing entails." While "extremely broad from a disciplinary perspective", the program is "tightly focused on specific applications." Issues covered in this initiative include "understanding humans as individuals, understanding humans as societies, computational models of behavior, social and cultural issues (diversity, culture, group dynamics, and technological change), economic impacts of IT, human-centered interfaces, human-centered applications, and human- centered systems."

The following is a list of HCC topics including the ones listed by the U.S. National Science Foundation in their HCC computing cluster (see [48]):

- Systems for problem-solving by people interacting in distributed environments. For example, in Internet-based information systems, in sensor-based information networks, and mobile and wearable information appliances;

- Multimedia and multimodal interfaces in which combinations of images and video, speech, text, graphics, gesture, touch, sound, etc. are used by people to communicate with one another;
- Intelligent interfaces and user modeling, information visualization, and adaptation of content to accommodate different display capabilities, modalities, bandwidth and latency;
- multi-agent systems that control and coordinate actions and solve complex problems in distributed environments in a wide variety of domains, such as ecommerce, medicine, or education;
- Models for effective computer-mediated human-human interaction under a variety of constraints, (e.g., video conferencing, collaboration across high vs. low bandwidth networks, etc.);
- Collaborative systems that enable knowledge-intensive and dynamic interactions for innovation and knowledge generation across organizational boundaries, national borders, and professional fields;
- Methods to support and enhance social interaction, including innovative ideas like social orthotics, affective computing, and experience capture;
- Social dynamics modeling and socially aware systems.

In terms of applications, the possibilities are endless if we wish to design and deploy computers using HCC methodologies. If we view computing as a large space with the human in the center, we can certainly identify some applications/fields that are closer to the human and some that are further away. Using an extreme example, packet switching is very important in communications, but its distance from a human is much larger, than, for instance, human-computer interaction. As computers become more pervasive, the areas that are closer to humans increase in number. If we view this historically, it is clear that computers are increasingly getting - physically, conceptually, and functionally - closer to humans (think of the first computers, those used in the 1950s and the current mobile devices), motivating the need for well-defined methodologies and philosophies for HCC.

Both HCC's scope and its possibilities are wide, but in our view, three factors form the core of HCC system and algorithm design processes:

- augmenting or taking into account individual human abilities and limitations,
- social and cultural awareness, and
- adaptability across individuals and specific situations.

If we consider these factors in the design of systems and algorithms, HCC applications should exhibit the following qualities:

- integrate input from different types of sensors and communicate through a combination of media as output,
- act according to the social and cultural context in which they are deployed, and
- be useful to diverse individuals in their daily life.

2.4 HCC, HCI, CSCW, User-centered Design, and Human Computation

Many important contributions to HCC have been made in fields such as HCI, computer-supported cooperative work (CSCW), user-centered design, cognitive psychology, sociology, anthropology, and others.

User-centered design can be characterized as a multistage problem-solving process that requires designers not only to analyze and foresee how users are likely to use an interface but also to test the validity of their assumptions with regard to user behavior in the real world [41]. Traditional design approaches often include heuristic evaluations and measurements of productivity and efficiency.

CSCW combines the understanding of the way people work in groups with the enabling technologies of computer networking and associated hardware, software, services, and techniques [25].

In contrast, HCC covers more than the traditional areas of usability engineering, HCI, and human factors, which are primarily concerned with user interfaces or user interaction. HCC "incorporates the learning, social, and cognitive sciences and intelligent systems areas more closely than traditional HCI" [20]. According to [26], compared to HCI, the shift in perspective with HCC is twofold:

- HCC is "conceived as a theme that is important for all computer-related research, not as a field which overlaps or is a sub-discipline of computer science";
- The HCC view acknowledges that "computing connotes both concrete technologies (that facilitates various tasks) and a major social and economic force."

Interactive evolutionary computation, an interesting technique first developed in the early 1990s and more recently known as human-based computation [22][3][2], also puts the human at the center, but in a different way. In traditional computation, a human provides a formalized problem description to a computer and receives a solution to interpret. In human-based computation, the roles are reversed: The computer asks a person or a large number of people to solve a problem, then collects, interprets, and integrates their solutions. In other words, the computer "asks" the human to do the work it cannot do. This is done as a way to overcome technical difficulties: Instead of trying to get computers to solve problems that are too complex, use human beings. In human computation, the human is helping the computer solve difficult problems, but in HCC, the computer helps humans maximize their abilities regardless of the situation - the human need not be in front of a computer performing a computing task when dealing with the computer.

Although HCC and human-based computation approach computing from two different perspectives, they both try to maximize the synergy between human abilities and computing resources. Work in human-based computation can therefore be of significant importance to HCC. On one hand, data collected through human-based computation systems can be valuable for developing machine-learning models. On the other hand, it can help us better understand human behavior and abilities, again of direct use in HCC algorithm development and system design.

HCC researchers also bring a diverse array of conceptual and research tools to traditional computing areas. HCC is not just about the interaction, the interface, or the design process. It is about knowledge, people, technology, and everything that ties them together. This includes how the technology is actually used and in what context. Furthermore, HCC involves both the creation of theoretical frameworks and the design and implementation of technical approaches and systems in many areas. Additionally, Dertouzos [16] points out that HCC goes well beyond user-friendly interfaces. This is because HCC uses five technologies in a synergistic way: natural interaction, automation, individualized information access, collaboration, and customization.

In considering the scope, application areas, and qualities of HCC systems, much can be learned from fields and disciplines related to HCC. As we will see in the next section, this leads us directly to focus on three core areas for Human-Centered Computing: production, analysis, and interaction.

3 Areas of Human-centered Computing

The distinctions between computing and multimedia computing are blurring very quickly. In spite of this, we can identify the main human-centered activities as follows [32]: *media production*, *annotation*,*organization*, *archival*, *retrieval*, *sharing*, *analysis*, and *communication*.

The above activities can in turn be clustered into three large activity areas: **production**, **analysis**, and **interaction**. These three areas are proposed here as a means to facilitate the discussion of the scope of HCC in the remainder of the paper. However, it must be evident that other ways of describing HCC are possible, and that the three proposed areas are interdependent in more than one way. Consider for example two typical multimedia scenarios which are representative to illustrate the HCC aspects.

In the first one, post-production techniques of non-edited home video normally use interaction (via manual composition of scenes), and increasingly rely on automatic analysis (e.g. shot boundary detection). In the second scenario, analysis techniques (e.g. automatic annotation of image collections) increasingly use metadata generated at production time (both automatic like time, date, camera type, etc. and human-produced via 'live' descriptions, commentaries, etc.), while performance can be clearly improved through various forms of interaction (e.g., via partial manual annotation, through active learning, etc.). In addition to this, further knowledge about humans could be introduced both at the individual level (in the first scenario, by adapting "film rules" to personalizing the algorithms, e.g. preferred modes of shooting a scene), and at the social level (in the second scenario, by using social context, provided for instance from a photo sharing website to improve annotation prediction for better indexing).

Each of the three main areas is discussed in the following sections, emphasizing social and cultural issues, and the integration of sensors and multiple media for system design, deployment, and access. Note that we use here interchangeably HCC and Human-center Multimedia (HCM) as the latter stresses the multimodal aspect of the former.

3.1 Multimedia Production

The first activity area in HCM is *multimedia production*, i.e. the human task of creating media (e.g., photographing, recording audio, combining, remixing, etc.). Although media can be produced automatically without human intervention once a system is set up (e.g., video from surveillance cameras), in HCM we are concerned with all aspects of media production which directly involve humans. Without a doubt, social and cultural differences result in differences in content at every level of the content production chain and at every level of the content itself (e.g., see discussions on the content pyramid in [32][33][17]). This occurs from low-level features (e.g., colors have strong cultural interpretations) to high-level semantics (e.g., consider the differences in communication styles between Japanese and American business people).

A good example of how cultural differences determine the characteristics of multimedia content is the TREC Video Retrieval Evaluation (TRECVID) set [47]. In news programs in some Middle Eastern countries there are mini-soap segments between news stories. Furthermore, the direction of text banners differs depending on language, and the structure of the news itself varies from country to country. The cultural differences are even greater in movies: colors, music, and all kinds of social and cultural signals convey the elements of a story (consider the differences between Bollywood and Hollywood movies in terms of colors, music, story structure, and so on).

Content is knowledge and vice versa: in HCM systems, cultural and social factors should ideally be (implicitly or explicitly) embedded in media at production time. In addition, HCM production systems should consider cultural differences and be designed according to the cultural context in which they will be deployed. As a simple example, a system for news editing in the Middle East might, by default, animate banners from left to right, or have special functions to distribute mini-soap segments across a newscast.

HCM production systems should also consider human abilities. For example, many of the systems for continuous recording of personal experiences [11] are meant to function as memory prosthesis, so one of their main goals is to record events with levels of detail (e.g., video, audio) that humans are incapable of recording (or rather, recalling). Interestingly, for these types of applications, integration of different types of sensors is also important, as is their adaptability, to each individual and particular context (a “user” of such system would probably not want everything to be recorded).

Unfortunately, in terms of culture, the majority of tools for content production follow a standard Western model, catering to a small percentage of the world’s population and ignoring the content gap (see the World Summit Award - http://www.wsis-award.org - which is an initiative to create an awareness of this gap) [8][38]. In terms of integration of multiple sensors, there has been some progress (e.g., digital cameras with GPS information or audio annotation, among others), but the issue of adaptability in content production systems has been seldom addressed. The result is that in spite of the tremendous growth in the availability of systems for media production, the field is in its infancy if we think of the population as a whole (relatively speaking, few people have access to computers, and out of those that do, even fewer can easily produce structured multimedia content with the current available tools).

3.2 Multimedia Analysis

A second activity area of great importance in HCM is automatic analysis of multimedia content. As described above, automatic analysis can be integrated in production systems (e.g., scene cut detection in video editing software). Interestingly, it can also alleviate some of the limitations in multimedia production because automatic analysis can be used to give content structure (e.g., by annotating it), increasing its accessibility. This has application in many Human-Centered areas (e.g., broadcast video, home video, data mining of social media, web search, etc.).

An interesting HCM application that has emerged in recent years is the automatic analysis of human activities and social behavior. Automatic analysis of social interaction finds a number of potentially relevant uses, from facilitating and enhancing human communication (on-line), to allowing for improved information access and retrieval (off-line), in the professional, entertainment, and personal domains.

Social interaction is inherently multimodal, and often recorded in multimedia form (e.g., video and information from other sensors). Unlike the traditional HCI view, which emphasizes communication between a person and a computer, the emphasis of an emerging body of research has shifted towards the study of computational models of human-to-human communication in natural situations. Such research has appeared in various communities under different names (social computing, socially-aware computing, computers-in-the-human-interaction-loop, etc.) [61]. Such interest has been boosted by the increasing capacity to acquire media with both fixed and mobile sensors and devices, and also to the ability to record and analyze largescale social activities through the internet (media sharing sites, blogs, etc). Social context can be provided through the understanding of patterns that emerge from human interaction at various temporal, spatial, and social scales, ranging from short-duration, face-to-face social signals and behaviors exchanged by peers or groups involving a few people, including interest, attraction [23], to mid-duration relations and roles that people play within groups, like influence and dominance [64], to group dynamics and trends that often emerge over extended periods of times, including degree of group membership, social network roles, group alliances, etc. [52].

There is no doubt of the importance of considering social and cultural differences in the design and application of algorithms and systems for multimedia analysis of human activities and social behavior. In turn, culture-specific knowledge should also be used in designing automatic analysis algorithms of multimedia in other domains in order to improve performance. In the news example from Section 3.1, an automatic technique designed for the US news style is likely to yield low performance when applied to a Middle Eastern newscast.

Augmenting or considering human abilities is also clearly beneficial because as argued earlier, there is a tight integration between the three activity areas we are considering, thus, what analysis algorithms are designed to do has a direct impact on how humans use multimedia data. The benefit of integrating multiple sensors is clear in the analysis of human activities (e.g., using input from RFID tags gives us information not easily attainable from video), as is the adaptability of HCM analysis systems to specific collections, needs, or particular tasks.

3.3 Multimedia Interaction

In addition to understanding the subjacent human tasks, the understanding of the multiplicity of forms that interaction can take is of particular importance for multimedia research within the HCM paradigm. In other words: it is paramount to understand both how humans interact with each other and why, so that we can build systems to facilitate such communication and so that people can interact with computers (or whatever devices embed them) in natural ways. We illustrate this point with three cases. In face-to-face communication, interaction is physically located and real-time. Concrete examples include professional settings like interviews, group meetings, and lectures, but also informal settings, including peer conversations, social gatherings, traveling, etc. The media produced in many of these situations might be in multiple modalities (voice, images, text, data from location, proximity, and other sensors), be potentially very rich, and often unedited (raw content). In a second case, live computer-mediated communication - ranging from the traditional teleconferencing and remote collaboration paradigms to emerging ubiquitous approaches based on wearable devices - is physically remote but remains real-time. In this case, the type of associated media will often be more limited or pre-filtered compared to the face-to-face case, due to bandwidth constraints. A final case corresponds to non-real time computer-mediated communication - including for instance SMS, mobile picture sharing, e-mail, blogging, media sharing sites, etc. - where, due to their own nature, media will often be edited, and interaction will potentially target larger, physically disjoint groups.

Unlike in traditional HCI applications (a single user facing a computer and interacting with it via a mouse or a keyboard), in the new applications (e.g., intelligent homes [1], remote collaboration, arts, etc.), interactions are not always explicit commands, and often involve multiple users. This is due in part to the remarkable progress in the last few years in computer processor speed, memory, and storage capabilities, matched by the availability of many new input and output devices that are making ubiquitous computing [63] a reality. Devices include phones, embedded systems, PDAs, laptops, wall size displays, and many others. The wide range of computing devices available, with differing computational power and input/output capabilities, means that the future of computing is likely to include novel ways of interaction and for the most part, that interaction is likely to be multimodal. Some of the modes of communication include gestures [50], speech [53], haptics [5], eye blinks [24], and many others. Glove mounted devices [7] and graspable user interfaces [18], for example, seem now ripe for exploration. Pointing devices with haptic feedback, eye tracking, and gaze detection [31] are also currently emerging. As in human-human communication, however, effective communication is likely to take place when different input devices are used in combination.

Given these trends, we view the interaction activity area of HCM as Multimodal interaction (see [34] and [35] for recent reviews). Clearly, one of the main goals of a HCC approach to interaction is to achieve natural interaction, not only with computers as we think of them today (i.e., machines on a desk), but rather with our environment, and with other people. Inevitably, this implies that we must consider culture because the way we generate signals and interpret symbols depends entirely on our cultural background. Multimedia systems should therefore use cultural cues during interaction [32] (such as a cartoon character bowing when a user initiates a transaction at an ATM). Although intuitively this makes sense, the majority of work in multimedia interaction assumes a one-size-fits-all model, in which the only difference between systems deployed in different parts of the world (or using different input data) is language. The spread of computing under the language-only difference model means people are expected to adapt to the technologies imposed arbitrarily using Western thought models. Clearly, these unfortunate trends are also due to social and economic factors, but as computing spreads beyond the desktop, researchers and developers are recognizing the importance of rethinking what we could call the "neutral culture syndrome" where it is erroneously believed that current computing systems are not culture specific.

In order to succeed, HCM interaction systems must be designed considering cultural differences and social context so that natural interaction can take place. This will inevitably mean that most HCM systems should embrace multimodal interaction, because multimodal systems open the doors to natural communication and to the possibility of adapting to particular users. Of course, integration of multiple sensors and adaptability are essential in HCM interaction. We describe some examples in the next section.

4 Applications

The range of application areas for HCC touches on many aspects of computing, and as computing becomes more ubiquitous, practically every aspect of interaction with objects, and the environment, as well as human-human interaction (e.g., remote collaboration, etc.) will make use of HCC techniques. In the following sections, we describe briefly specific application areas in which interesting progress has been made (for more details see [35]).

4.1 Human Spaces

Computing is expanding beyond the desktop, integrating with everyday objects in a variety of scenarios. As our discussions show, this implies that the model of user interface in which a person sits in front of a computer is no longer the only model. One of the implications of this is that the actions or events to be recognized by the "interface" are not necessarily explicit commands. In smart conference room applications, for instance, multimodal analysis has been applied mostly for video indexing [45] (see [52] for a social analysis application). Although such approaches are not meant to be used in real-time, they are useful in investigating how multiple modalities can be fused in interpreting communication. It is easy to foresee applications in which "smart meeting rooms" actually react to multimodal actions in the same way intelligent homes should [1]. Projects in the video domain include MVIEWS [12], a system for annotating, indexing, extracting, and disseminating information from video streams for surveillance and intelligence applications. An analyst watching one or more live video feeds is able to use pen and voice to annotate the events taking place. The annotation streams are indexed by speech and gesture recognition technologies for later retrieval, and can be quickly scanned using a timeline interface, then played back during review of the film. Pen and speech can also be used to command various aspects of the system, with multimodal utterances such as "Track this" or "If any object enters this area, notify me immediately."

4.2 Ubiquitous Devices

The recent drop in costs of hardware has led to an explosion in the availability of mobile computing devices. One of the major challenges is that while devices such as PDAs and mobile phones have become smaller and more powerful, there has been little progress in developing effective interfaces to access the increased computational and media resources available in such devices. mobile devices, as well as wearable devices, constitute a very important area of opportunity for research in HCC because natural interaction with such devices can be crucial in overcoming the limitations of current interfaces. Several researchers have recognized this, and many projects exist on mobile and wearable HCC [9][21][51].

4.3 Users with Disabilities

People with disabilities can benefit greatly from HCC technologies [42]. Various authors have proposed approaches for smart wheel-chair systems which integrate different types of sensors. The authors of [55] introduce a system for presenting digital pictures non-visually (multimodal output), and the techniques in [24] can be used for interaction using only eye blinks and eye brow movements. Some of the approaches in other application areas (e.g., [9][60]) could also be beneficial for people with disabilities.

4.4 Public and Private Spaces

In this category we place applications implemented to access devices used in public or private spaces. One example of implementation in public spaces is the use of HCC in information kiosks [40][54]. These are challenging applications for natural multimodal interaction: the kiosks are often intended to be used by a wide audience, thus there may be few assumptions about the types of users of the system. On the other hand, there are applications in private spaces. One interesting area is that of implementation in vehicles [39][58][59]. This is an interesting application area due to the constraints: since the driver must focus on the driving task, traditional interfaces (e.g., GUIs) are not so suitable. Thus, it is an important area of opportunity for HCC research, particularly because depending on the particular deployment, vehicle interfaces can be considered safety-critical.

4.5 Virtual Environments

Virtual and augmented reality has been a very active research area at the crossroads of computer graphics, computer vision, and human-computer interaction. One of the major difficulties of VR systems is the interaction component, and many researchers are currently exploring the use of interaction analysis techniques to enhance the user experience. One reason this is very attractive in VR environments is that it helps disambiguate communication between users and machines (in some cases virtual characters, the virtual environment, or even other users represented by virtual characters [46]).

4.6 Art

Perhaps one of the most exciting application areas of HCC is art. Vision techniques can be used to allow audience participation [44] and influence a performance. In

[62], the authors use multiple modalities (video, audio, pressure sensors) to output different "emotional states" for Ada, an intelligent space that responds to multimodal input from its visitors. In [43], a wearable camera pointing at the wearer's mouth interprets mouth gestures to generate MIDI sounds (so a musician can play other instruments while generating sounds by moving his mouth). In [49], limb movements are tracked to generate music. HCC can also be used in museums to augment exhibitions [49].

4.7 Other

Other applications include education, remote collaboration, entertainment, robotics, surveillance, or biometrics. HCC can also play an important role in safety-critical applications (e.g., medicine, military, etc.) and in situations in which a lot of information from multiple sources has to be viewed in short periods of time. A good example of this is crisis management.

5 Integrating HCC into a Real (Human) World

Human-Centered Computing systems and applications should ultimately be integrated in a world that is complex and rapidly evolving. For instance, computing is migrating from the desktop, at the same time as the span of users is expanding dramatically to include people who would not normally access computers. This is important because although in industrialized nations almost everyone has a computer, a small percentage of the world's population owns a multimedia device (millions still do not have phones). The future of multimedia, therefore, lies outside the desktop, and multimedia will become the main access mechanism to information and services across the globe. Integration of modalities and media, of access mechanisms, and of resources constitute three key-issues for the creation of future HCC systems. We discuss each of these issues in the following subsections (for more details see [32]).

5.1 Integrating Modalities and Media

Despite great efforts in the multimedia research community, integrating multiple media (in production, analysis, and interaction) is still in its infancy. Our ability to communicate and interpret meanings depends entirely on how multiple media are combined (such as body pose, gestures, tone of voice, and choice of words), but most research on multimedia focuses on a single medium model. In the past, interaction concerns have been left to researchers in HCI and the scope of work on interaction within the multimedia community has focused mainly on image and video browsing. Multimedia, however, includes many types of media and, as evidenced by many projects developed in the arts, multimedia content is no longer limited to audiovisual materials. Thus, we see interaction with multimedia data not just as an HCI problem, but as a multimedia problem. Our ability to interact with a multimedia collection depends on how the collection is indexed, so there is a tight integration between analysis and interaction. In fact, in many multimedia systems we actually interact with multimedia information and want to do it in a multimodal way.

Two major research challenges are modeling the integration of multiple media in analysis, production, and multimodal interaction. Statistical techniques for modeling are a promising approach for certain types of problems. For instance, Dynamic Bayesian Networks have been successfully applied in a wide range of problems that have a time component, while sensor fusion and classifier integration in the artificial intelligence community have also been active areas of research. In terms of content production, we do not have a good understanding of the human interpretation of the messages that a system sends when multiple media are fused and there is much we can learn from the arts and communication psychologists.

Because of this lack of integration, existing approaches suit only a small subset of the problems and more research is needed, not only on the technical side, but also on understanding how humans actually fuse information for communication. This means making stronger links between fields like neuroscience, cognitive science, and multimedia development. For instance, exploring the application of Bayesian frameworks to integration [4], investigating different modality fusion hypothesis [15] (discontinuity, appropriateness, information reliability, directed attention, and so on), or investigating stages of sensory integration [56] can potentially give us new insights that lead to new technical approaches.

Without theoretical frameworks on integrating multiple sensors and media, we are likely to continue working on each modality separately and ignoring the integration problem, which should be at the core of multimedia research.

5.2 Integrating Access

Everyone seems to own a mobile device. As a consequence, there is a new wave of portable computing, where a cell phone is a fully functional computer that we can use to communicate, record, and access a wealth of information (such as location-based, images, video, personal finances, and contacts). Although important progress has been made, particularly in ambient intelligence applications [1] and in the use of metadata from mobile devices [6][14], much work needs to be done and one of the technical challenges is dealing with large amounts of information effectively in real time. Developing effective interaction techniques for small devices is one of our biggest challenges because strong physical limitations are in place. In the past, we assumed the desktop screen was the only output channel, so advances in mobile devices are completely redefining multimedia applications. But mobile devices are used for the entire range of human activities: production, annotation, organization, retrieval, sharing, communication, and content analysis.

5.3 Integrating Resources

While mobile phone sales are breaking all records, it is increasingly common for people to share computational resources across time and space. Public multimedia devices are becoming increasingly common. In addition, it is important to recognize that Ů particularly in developing countries Ů sharing of resources is often the only option. Many projects for sharing community resources exist, particularly for rural areas, in education and other important activities. One of the main technical research challenges here is constructing scalable methods of multimodal interaction that can quickly adapt to different types of users, irrespective of their particular communication abilities.

The technical challenges in these two cases seem significantly different: mobile devices should be personalized, while public systems should be general enough to be effective for many different kinds of users. Interestingly, however, they both fall under the umbrella of ubiquitous multimedia access: the ability to access information anywhere, anytime, on any device. Clearly, for these systems to succeed we need to consider cultural factors (for example, text messaging is widespread in Japan, but less popular in the US), integration of multiple sensors, and multimodal interaction techniques.

In either case, it is clear that new access paradigms will dominate the future of computing and ubiquitous multimedia will play a major role. Ubiquitous multimedia systems are the key in letting everyone access a wide range of resources critical to economic and social development.

6 Research Agenda for HCC

To summarize the major points that we have presented so far, human-centered computing systems should be multimodal (inputs and outputs in more than one modality or communication channel), they must be proactive (understand cultural and social contexts and respond accordingly), and be easily accessible outside the desktop to a wide range of users (i.e., adaptable) (see Section 2.3 and [32]).

A human-centered approach to multimedia will consider how humans understand and interpret multimedia signals (feature, cognitive, and affective levels), and how humans interact naturally (the cultural and social contexts as well as personal factors such as emotion, mood, attitude, and attention). Inevitably, this means considering some of the work in fields such as neuroscience, psychology, cognitive science, and others, and incorporating what is known in those fields within computational frameworks that integrate different media.

Research on machine learning integrated with domain knowledge, automatic analysis of social networks, data mining, sensor fusion research, and multimodal interaction will play a special role [52]. Further research into quantifying human-related knowledge is necessary, which means developing new theories (and mathematical models) of multimedia integration at multiple levels. We believe that a research agenda on HCC will involve the following nonexhaustive list of goals [36][37]:

- New human-centered methodologies for the design of models and algorithms and the development of systems in each of the areas discussed previously.
- Focused research on the integration of multiple sensors, media, and human sciences that have people as the central point.
- New interdisciplinary academic and industrial programs, initiatives, and meeting opportunities.
- Discussions on the impact of multimedia technology that include the social, economic, and cultural contexts in which such technology is or might be deployed.
- Research data that reflect the human-centered approach, e.g., data collected from real social situations, data that is rich - multisensorial and culturally diverse.
- Common computing resources on HCM (e.g. software tools and platforms).
- Evaluation metrics for theories, design processes, implementations, and systems from a human-centered perspective; and
- Methodologies for privacy protection and the consideration of ethical and cultural issues.

Human-centered approaches have been the concern of several disciplines [19] but, as pointed out in Section 1, they have been often undertaken in separate fields. The challenges and opportunities in the field of multimedia are great not only because so many of the activities in multimedia are human-centered, but also because multimedia data itself is used to record and convey human activities and experiences. It is only natural, therefore, for the field to converge in this direction and play a key role in the transformation of technology to truly support people's activities.

7 Conclusions

In this chapter, we gave an overview of HCC from a several perspectives. We described the three main areas of Human-Centered Computing emphasizing social and cultural issues, and the integration of sensors and multiple media for system design, deployment, and access. A research agenda for HCC (see also [32][37]) was presented.

Many technical challenges lie ahead and in some areas progress has been slow. With the cost of hardware continuing to drop and the increase in computational power, however, there have been many recent efforts to use HCC technology in entirely new ways. One particular area of interest is new media art. Many universities around the world are creating new joint art and computer science programs in which technical researchers and artists create art that combines new technical approaches or novel uses of existing technology with artistic concepts. In many new media art projects, technical novelty is introduced while many HCC issues are considered: cultural and social context, integration of sensors, migration outside the desktop, and access.

Technical researchers need not venture into the arts to develop human-centered computing systems. In fact, in recent years many human-centered multimedia applications have been developed within the multimedia domain (such as smart homes and offices, medical informatics, computerguided surgery, education, multimedia for visualization in biomedical applications, education, and so on). However, more efforts are needed and the realization that multimedia research, except in specific applications, is meaningless if the user is not the starting point. The question is whether multimedia research will drive computing (with all its social impacts) in synergy with human needs, or it will be driven by technical developments alone.

References

[1] Aarts E (2004) Ambient intelligence: A multimedia perspective. IEEE Multimedia 11(1):12–19
[2] von Ahn L (2006) Games with a purpose. IEEE Computer pp 96–98
[3] von Ahn L, Dabbish L (2008) General techniques for designing games with a purpose. Communications of the ACM pp 58–67
[4] Andersen T, Tiippana K, Sams M (2004) Factors influencing audiovisual fission and fusion illusions. Cognitive Brain Research 21:301–308
[5] Benali-Khoudja M, Hafez M, Alexandre JM, Kheddar A (2004) Tactile interfaces: A state-of-the-art survey. In: Int. Symposium on Robotics
[6] Boll S (2005) Image and video retrieval from a user-centered mobile multimedia perspective. In: Int'l Conf. Image and Video Retrieval
[7] Borst C, Volz R (2005) Evaluation of a haptic mixed reality system for interactions with a virtual control panel. Presence: Teleoperators and Virtual Environments 14(6)
[8] Brewer E (2005) The case for technology for developing regions. IEEE Computer 38(6):25–38
[9] Brewster S, Lumsden J, Bell M, Hall M, Tasker S (2003) Multimodal 'eyes-free' interaction techniques for wearable devices. In: ACM CHI
[10] Canny J (2001) Human-center computing. Tech. rep., UC Berkeley HCC Retreat
[11] CARPE (2005) ACM workshop on capture, archival, and retrieval of personal experiences (carpe)
[12] Cheyer A, Julia L (1998) MVIEWS: Multimodal tools for the video analyst. In: Intelligent User Interfaces

[13] Clancey W (1997) Situated Cognition: On Human Knowledge and Computer Representations. Cambridge University Press
[14] Davis M, Sarvas R (2004) Mobile media metadata for mobile imaging. In: Int'l Conf. Multimedia and Expo
[15] Deneve S, Pouget A (2004) Bayesian multisensory integration and cross-modal spatial links. J Physiology 98(1-3):249–258
[16] Dertouzos M (2001) The Unfinished Revolution: Human-centered Computers and What They Can Do for Us. HarperCollins
[17] Dimitrova N (2004) Context and memory in multimedia content analysis. IEEE Multimedia 11(3):7–11
[18] Fitzmaurice G, Ishii H, Buxton W (1995) Bricks: Laying the foundations for graspable user interfaces. In: ACM CHI
[19] Flanagan J, Huang T, Jones P, Kasif S (1997) Human-centered systems: Information, interactivity, and intelligence. Tech. rep., NSF
[20] Foley J (2008) HCC educational digital library. Tech. rep., http://hcc.cc.gatech.edu
[21] Fritz G, Seifert C, Luley P, Paletta L, Almer A (2004) Mobile vision for ambient learning in urban environments in urban environments. In: Int. Conf. on Mobile Learning
[22] Gentry C, Ramzan Z, Stubblebine S (2005) Secure distributed human computation. In: ACM Conf. on Electronic Commerce, pp 328–332
[23] Gladwell M (2005) Blink. Little Brown and Co.
[24] Grauman K, Betke M, Lombardi J, Gips J, Bradski G (2003) Communication via eye blinks and eyebrow raises: Videobased human-computer interfaces. Universal Access in Inf Society 2(4):359–373
[25] Grudin J (1994) Computer-supported cooperative work: Its history and participation. IEEE Computer pp 19–26
[26] HCC (2008) Human-Centered Computing: An interdisciplinary consortium at UC Berkeley . URL `http://www.cs.berkeley.edu/jfc/hcc/`
[27] HCC-PhD (2008) PhD Human Centered Computing. URL `http://www.cc.gatech.edu/education/grad/phd-hcc`
[28] Hoffman R (2004) Human-centered computing principles for advanced decision architectures. Tech. rep., Army Research Laboratory
[29] Hoffman R, Feltovich P, Ford K, Woods D, Klein G, Feltovich A (2002) A rose by any other name ... would probably be given an acronym. IEEE Intelligent Systems 17(4):72–80
[30] IHMC (2008) Institute of Human and Machine Cognition. URL `http://www.ihmc.us`
[31] Jacob R (2005) The use of eye movements in human-computer interactions techniques: What you look at is what you get. ACM Transactions Information Systems 9(3):152–169
[32] Jaimes A (2006) Human-centered multimedia: Culture, deployment, and access. IEEE Multimedia 13(1):12–19
[33] Jaimes A, Chang SF (2000) A conceptual framework for indexing visual information at multiple levels. In: SPIE Internet Imaging

[34] Jaimes A, Sebe N (2005) Multimodal hci: A survey. In: IEEE Int'l Workshop on Human-Computer Interaction
[35] Jaimes A, Sebe N (2007) Multimodal human computer interaction: A survey. Computer Vision and Image Understanding 108(1-2):116–134
[36] Jaimes A, Sebe N, Gatica-Perez D (2002) Human-centered computing: A multimedia perspective. In: ACM Multimedia, pp 855–864
[37] Jaimes A, Gatica-Perez D, Sebe N, Huang T (2007) Human-centered computing: Toward a human revolution. IEEE Computer 40(5):30–34
[38] Jain R (2003) Folk computing. Comm of the ACM 46(4):27–29
[39] Ji Q, Yang X (2002) Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. Real-Time Imaging 8:357–377
[40] Johnston M, Bangalore S (2002) Multimodal applications from mobile to kiosk. In: W3C Workshop on Multimodal Interaction
[41] Karat J, Karat C (2003) The evolution of user-centered focus in the human-computer interaction field. IBM Systems 42(2):532–541
[42] Kuno Y, Shimada N, Shirai Y (2003) Look where you're going: A robotic wheelchair based on the integration of human and environmental observations. IEEE Robotics and Automation 10(1):26–34
[43] Lyons M, Haehnel M, Tetsutani N (2003) Designing, playing, and performing, with a vision-based mouth interface. In: Conf. on New Interfaces for Musical Expression
[44] Maynes-Aminzade D, Pausch R, Seitz S (2002) Techniques for interactive audience participation. In: ICMI
[45] McCowan I, Gatica-Perez D, Bengio S, Lathoud G, Barnard M, Zhang D (2005) Automatic analysis of multimodal group actions in meetings. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(3):305–317
[46] Nijholt A, Heylen D (2002) Multimodal communication in inhabited virtual environments. Int J of Speech Technology 5:343–354
[47] NIST-TRECVID (2008) TREC Video Retrieval Evaluation. URL http://www-nlpir.nist.gov/projects/trecvid
[48] NSF-HCC (2008) NSF - Human Centered Computing Program. URL http://www.nsf.gov/cise/iis/hcc_pgm.jsp
[49] Paradiso J, Sparacino F (1997) Optical tracking for music and dance performance. In: Optical 3-D Measurement Techniques
[50] Pavlovic V, Sharma R, Huang T (1997) Visual interpretation of hand gestures for human-computer interaction: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7):677–695
[51] Pelz J (2004) Portable eye-tracking in natural behavior. J of Vision 4
[52] Pentland A (2005) Socially aware computation and communication. IEEE Computer 38(3):33–40
[53] Potamianos G, Neti C, Luettin J, Matthews I (2004) Audiovisual automatic speech recognition: An overview. In: Issues in Visual and Audio-Visual Speech Processing, MIT Press
[54] Rehg J, Loughlin M, Waters K (1997) Vision for a smart kiosk. In: CVPR

[55] Roth P, Pun T (2003) Design and evaluation of a multimodal system for the non-visual exploration of digital pictures. In: INTERACT
[56] Schroeder C, Foxe J (2005) Multisensory contributions to low-level, 'unisensory' processing. Current Opinion in Neurobiology 15:454–458
[57] Shneiderman B (2002) Leonardo's Laptop: Human Needs and the New Computing Technologies. MIT Press
[58] Smith P, Shah M, Lobo N (2003) Determining driver visual attention with one camera. IEEE Transactions on Intelligent Transportation Systems 4(4)
[59] Trivedi M, Gandhi T, McCall J (2007) Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety. IEEE Transactions on Intelligent Transportation Systems pp 108–120
[60] Valenti R, Gevers T (2008) Accurate eye center location and tracking using isophote curvature. In: IEEE Conf. Computer Vision and Pattern Recogn.
[61] Vinciarelli A, Pantic M, Bourlard H, Pentland A (2008) Social signal processing: State-of-the-art and future perspectives of an emerging domain. In: ACM Multimedia, pp 1061–1070
[62] Wassermann K, Eng K, Verschure P, Manzolli J (2003) Live soundscape composition based on synthetic emotions. IEEE Multimedia 10(4)
[63] Weiser M (1993) Some computer science issues in ubiquitous computing. Communications of the ACM 36(7):74–83
[64] Zhang D, Gatica-Perez D, Bengio S, Roy D (2005) Learning influence among interacting Markov chains. In: NIPS