Related Work Description and Some Notes

Yizheng Gu
Tsinghua University
Beijing, China
guyz17@mail.tsinghua.edu.cn

2nd Author Name
Affiliation
City, Country
e-mail address

3rd Author Name
Affiliation
City, Country
e-mail address

ABSTRACT
While 3D Tele-immersion techniques develop rapidly, the effects of transmission quality on its users remain unexplored. Since the transmission quality such as delay, frame rate, resolution and cost are trade-off nowadays, we argue that it is important to study the user experience in 3D Tele-immersion systems.

Author Keywords
Telepresence, Delay, Network Performance, QOE.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; See http://acm.org/about/class/1998 for the full list of ACM classifiers. This section is required.

INTRODUCTION
注：以下还不是 introduction，只是我自己做的一些笔记。

看起来，3D Tele-immersion 相比于其它更简单的技术而言，其优势并不是我们想象中那么理所当然，也许需要在 introduction 中讨论一下。

我似乎忽略了 FOV 等因素。这些因素不属于 transmission quality trade-off，但是对人的认知是重要的。

我们在实验上的贡献不多说了。我们在技术上的贡献：1、我们尝试对齐并 merge 两个场景中相同的物体，以提供 shared objects in the virtual scene；2、我们提供一个可操作的交互过程，来校准两个场景，裁剪无用的背景，凸显公共部分和人物。3、我们尝试了并对比了许多商用硬件，发现 RealSense 在深度摄像头中最优，发现 Gtx1080Ti 可以实现多少多少的实时性，而 HTC Vive 则免去已渲染图像无线传输的不便，我们还公布了经过深度 GPU 优化的代码；4、这套系统除了网络传输以外，使用了可以在实验室外也能获得的设备，是 practicable 的，而网络传输部分只需要了 2D 视频

流，经检验，带宽为多少多少。

RELATED WORK
Related work 的思路暂定如下：（1）首先介绍基于 Reconstruction 的 3DTI 技术（Tele-immersion）有哪些，突出其好处，表明基于 Reconstruction 的 3DTI 将会成为下一代的远程通信技术；（2）在之前的工作中，存在对基于 robotic 或者 presed-3D 远程呈现的研究，也有对 Tele-immersion 中一些小问题进行的小型 user study，但对 Tele-immersion 中最基本的问题，如延迟对用户的影响，还不存在；（3）为了开展这份工作，我们参考了 Telephone 和 Video conference 中的测量方法，这里总结一下其中的 measurement 方法。或者写 QoE 的相关研究。

在 Related work 的写作过程中，第一部分将参考 Holoportation，第二部分待定，第三部分将参考 [General Recommendation on the transmission quality for an entire international telephone connection]。

3DTI technique
In early works, researchers use an array of cameras to construct the virtual scene from physical world.

[Virtual space teleconferencing using a sea of cameras. 1994]

[Virtualized reality: Constructing virtual worlds from real scenes. 1997]

[TELEPORT. 1999] the early immersive telepresence system.

Then, researchers start to improve the capturing of 3D model in real time.

[3D Tele-Collaboration Over Internet2. 2002] researchers assembled the best of available hardware and software technologies to present a 3D tele-immersion with point cloud.

[blue-c. 2003] is an early telepresence system.

[Real World Video Avatar. 2004] transmit and present the figure of a person at a remote location in real time.

[Immersive 3D Environment for remote collaboration. 2008] presented an immersive telepresence system for remote collaboration and training of physical activities.

[Multicamera Real-Time 3D Modeling for Telepresence and remote collaboration. 2010] This system allows to retrieve in real-time a 3D mesh of the observed scene.

[Encumbrance-Free Telepresence System. 2011] presented a 3D telepresence system which is affordable and reproducible, offering the opportunity to deliver it beyond the researcher's lab.

[KinectFusion. 2011] present a system for real-time mapping of arbitrary indoor scenes, using only a moving low-cost depth camera.

[Real-Time Volumetric 3D capture. 2012] described a 3D acquisition system, which is the most similar to our system. Our system is designed based on this previous work, the kernel algorithms of which are Tsdf Volume and Marching Cubes.

[A Volumetric Method for Building Complex. 1996] This paper describes the algorithm of creating the 3D model from range image.

[Marching Cubes. 1987] is a high resolution 3D surface construction algorithm based on TSDF Volume.

Recently, state-of-the art techniques achieve higher rendering quality.

[Fusion4D. 2016] contributed a pipeline for live high-quality reconstructions in real-time.

[Holoportation. 2016] for the first time presents an end-to-end system for augmented and virtual reality telepresence. This paper contributes as describing the tele-immersion pipeline in full.

User Study in 3D Telepresence
There are basically three types of techniques in 3D telepresence: Pseudo-3D, Robotic, Reconstruction. The impacts on users were well study in both robotic and embodiment, while the users' feeling in 3d reconstruction based telepresence remained underexplored.

Pseduo-3D
[ImmerseBoard]

[Towards Next Generation 3D Teleconferencing Systems]

[Evaluation of Factors Affecting Distance Perception] identify the most important factors that influence decision making and accuracy of distance perception with an immersive large screen display. The paper presents guidelines for setting up this kind of system.

Robotic
[The Future of Robotic Telepresence]

[Racetime: Telepresence Racing Game with]

[Movement Matters]

[My Student is a Robot]

[Robotic Telepresence at Scale]

[To Beam or Not to Beam]

[JackIn Head] JackIn Head allows remote user to see what the "Body user" can see. "Body user" wears an omnidirectional camera to transmit video to remote user. This paper contributes as an image stabilization technique and a study to this system.

[Augmented reality-aided telepresence system for robot] investigates the use of an immersive telepresence system for industrial robotics. Experiment shows the practicability and effectiveness.

[ChameleonMask: Embodied Physical and Social] uses a real human as a surrogate for the remote user.

[Can you see me now?] study the effects of FOV on users with robotic telepresence systems.

[Bodies in Motion] Mobility significantly increased the user's feelings of presence but decrease task performance.

[Communication Behavior in Embodied Virtual Reality] Maybe it is not robotic, but embodiment.

Reconstruction
[MirageTable] provides a one-to-one 3D telepresence experience by using Kinect and projector. Specially this system allows freehand physical interaction with virtual world.

[Immersive Group-to-Group Telepresence] This work built the first telepresence system which provides a shared space in immersive virtual reality environment. This paper and [Adapting standard video codecs for depth streaming] mention that they do not focus on the compression and transmission bandwidth.

[Room2Room] MSR proposed Room2Room system, which allows room to room naked eye 3d telepresence system with only a Kinect and a projector. They did a study and shows that it is significantly better in completion time and presence compared to Skype.

QoE in Telecommunication

Telephone
[General Recommendation on the transmission quality for an entire international telephone connection] For telecommunications, a latency which is longer than 400ms is unacceptable. Currently, 150ms is used as a rule of thumb, a value drawn by telecommunication research. The most applications would be acceptable within 150ms. For more immersive telecommunications such as 3D conferencing, a latency of 100ms would be better.

[The E-model] This model can help ensure that users will be satisfied with end-to-end transmission performance, avoiding over-engineering of networks. It is a tool for assessing the combines effects of variations in several transmission parameters the affect the conversational quality. The output is a scalar rating of transmission quality.

[Objective and Subjective assessment of telecommunication quality] In the annex of this paper it describes a user study to evaluate subjective feedback of telephone call in details.

Video Conference

In 2D telepresence, huge amount of works related to users' feeling has been done.

[Are We in sync?] In synchronization of watching video, 500ms is acceptable with video chat, and 2s is for text chat.

[Beyond Talking Heads] This paper in CSCW 2016 conducted a survey with members of distributed teams to investigate how they host distributed meetings. Nowadays most distributed meetings are hosted with video-based conference. The results suggest that tools and approaches are inadequate for meeting scenarios that require participants to not only converse but also to share different types of multimedia content. Note: 41% video-conference involve 3+ locations.

[Taking Notes or Playing Games?] presents a detailed examination of factors that affect perceptions of multitasking in video conferencing. It proposes a conceptual model for the design of video communication tools.

[Accuracy of Deictic Gestures to Support Telepresence] presents a controlled study to assess the accuracy when user showing a shored object on a wall-sized display. Eye gaze is more important than hand gesture.

[Remote handshaking] touch enhances video-mediated telepresence which partner's action needs to occur but should be invisible in the video.

Quality of Experience

[Augmented Sport. 2013] The Rational Part of UX in telepresence. *Psychology flow*: Concentration, Enjoyment and Sense of control. *Telepresence*: the sense of being or the sense of presence. *Technology Acceptance*: perceived usefulness and perceived easiness.

The Experiential Part of UX. *Social Presence*: social context, online communication and interactivity. *Social Emotion* are emotions that require the representation of the mental states of other people. *Emotional Response*.

This paper also pointed out the special different constructs in 3D LIVE UX and QoE model.

[Understanding How Network Performance Affects. 2014] Much research has been done to support remote collaboration on physical tasks, however, less attention has been paid to investigating how network performance can affect user experience and task performance. This paper examined 5 network scenarios and f metrics including *delay*, *jitter*, *bandwidth* and *packet loss*, and find that packet loss is the most significant in their system.

Delay: up to 400ms for telephone quality speech, maximum 100ms to 500ms for applications with real time sound transmission.

Jitter: the jitter for videoconferencing applications should not exceed 400ms.

QoE questions: Time taken for the task? Repeated instructions? Video/Audio/Overall quality rating?

There is a table of QoS metrics with Ethernet/Satellite/3G/Wi-Fi/Fibre.

[Requirements of Network Applications on the Internet. 2004] Real time sound is the most sensitive data transmission type to jitter.

[Definitions of terms related to quality service. 2008] QoS is defined by the ITU as "Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service".

QoE is defined as "the overall acceptability of an application or service, as perceived subjectively by the end-user".

[Interactivity Patterns in the QoE in Multi-party Video-mediated Conversations with delay. 2014] This paper took not only system factors but also context and user factors into account to investigate the QoE in Multi-party (more than 2) video-mediated conversations. The author suggested that user activity should be considered as a prime adaptation and optimization parameter.

When connecting two computers between New York and Hong Kong, the round delay is up to 776ms for Google+, and 1467ms for Skype.

[Qualinet White Paper on Definition of Quality of Experience. 2012] [A QoE Testbed for Socially-aware Video-mediated Group Communication] QoE is shaped mainly by the system, the user and the context.

[Stevens' Power Law in 3D Tele-immersion. 2014] Given a stimulus A and its responding quantified perception B, Stevens' Power Law indicates that $B = k * A \wedge p$. This paper found that this law has no significance of packet loss in 3D Tele-immersion.

[Quality of Experience in Immersive Multimedia. 2015] A clear distinction between low and high IL were observed, whereas the differences between middle and high IL were not significant.

[QoE Modelling for VP9 and H.265 Videos. 2015] This paper introduced novel predictions such as the correlations between video solution etc., and achieve a goodness-of-fit to the subjective assessment data (R>83%).

[VQMT: Video Quality Measurement Tool. 2013] as title say, PSNR, SSIM, etc.

[Exploring QoE for Power Efficiency. 2015] A field study on mobile videos with LCD displays.

FUTURE WORK

Beside delay, there are some other network performance can affect user experience, such as bandwidth (relating to rendering quality), jitter and packet loss. In the future, we will first figure out how should we assign the transmission data when the bandwidth is fixed, i.e., the resolution of depth image and color image flow. Then, we will further investigate jitter and packet loss and try to find out the solutions. Our final target is to totally understand how network performance affects user experience of 3D Tele-Immersion, and give a guideline for its engineering.

REFERENCES

TYPESET TEXT

The styles contained in this document have been modified from the default styles to reflect ACM formatting conventions. For example, content paragraphs like this one are formatted using the `Normal` style.

Title and Authors

Your paper's title, authors, and affiliations should run across the full width of the page in a single column 7 in (17.8 cm) wide. The title should be in Helvetica or Arial 18-point bold (the `Title` style in this document). Authors' names should be in Times New Roman or Times Roman

12-point bold (`Author Name` style), and affiliations in the font as 12-point regular (`Author Affiliation` style).

To position names and addresses, use a single-row table with invisible borders, as in this document. Alternatively, if only one address is needed, use a centered tab stop to center all name and address text on the page; for two addresses, use two centered tab stops, and so on. For more than three authors, you may have to place some address information in a footnote, or in a named section at the end of your paper. Leave one 10-point line of white space below the last line of affiliations.

Adding and Removing Author Fields Above

If you need a 4th block for an additional author with a different affiliation, click on the table, click "layout" under "table tools", and click "insert left" or "insert right". Then right-click anywhere on the table, click "Table Properties", and within the "Table" tab, click the tab box for "Preferred Width" and enter "7".

If you need author blocks for only 1 or 2 authors, you should remove one column from the table. Right-click in the unwanted cell, click "Delete Cell", click "Delete entire column", then click "OK". Repeat if necessary.

Abstract and Keywords

Every submission should begin with an abstract of about 150 words, followed by a set of keywords. The abstract and keywords should be placed in the left column of the first page under the left half of the title. The abstract should be a concise statement of the problem, approach, and conclusions of the work described. It should clearly state the paper's contribution to the field of HCI.

Normal or Body Text

Please use 10-point Times New Roman or Times Roman font (the `Normal` style) or, if this is unavailable, another proportional font with serifs, as close as possible in appearance to Times New Roman 10-point. Other than Helvetica or Arial headings, please use sans-serif or non-proportional fonts only for special purposes, such as source code text.

First Page Copyright Notice

This sample word document has the correct ACM SIGCHI copyright notice in place (see page 1, bottom of column 1). Accepted papers will be distributed in the conference publications. They will also be placed in the ACM Digital Library, where they will remain accessible to thousands of researchers and practitioners worldwide. ACM's copyright and permissions policy is here:
http://acm.org/publications/policies/copyright_policy

Subsequent Pages

On pages beyond the first, start at the top of the page and continue in double-column format. The two columns on the last page should be of approximately equal length.

Figure 1. Use high-resolution images, 300+ dpi, legible if printed in color or black-and-white. Number all figures and include captions below, using Insert, Caption.

References and Citations

Use a numbered list of references at the end of the article, ordered alphabetically by last name of first author, and referenced by numbers in brackets [1,3,4].

Objects	Caption – pre-2002	Caption – 2003 and afterwards
Tables	Above	Below
Figures	Below	Below

Table 1. Table captions should be placed below the table. We recommend table lines be 1 point, 25% black. Minimize use of unnecessary table lines.

Your references should be published materials accessible to the public. Internal technical reports may be cited only if they are easily accessible (i.e., you provide the address for obtaining the report within your citation) and may be obtained by any reader for a nominal fee. Proprietary information may not be cited. Private communications should be acknowledged in the main text, not referenced (e.g., "[Borriello, personal communication]").

References should be in ACM citation format: http://acm.org/publications/submissions/latex_style. This includes citations to internet resources [1,4,8,1] according to ACM format, although it is often appropriate to include URLs directly in the text, as above.

SECTIONS

The heading of a section should be in Arial 9-point bold, all in capitals (`Heading 1` style). Sections should not be numbered.

Subsections

Headings of subsections should be in Arial 9-point bold with initial letters capitalized (`Heading 2` style). For sub-sections and sub-subsections, a word like *the* or *of* is not capitalized unless it is the first word of the heading.

Sub-subsections

Headings for sub-subsections should be in Arial 9-point italic with initial letters capitalized (`Heading 3` style).

FIGURES/CAPTIONS

Place figures and tables at the top or bottom of the appropriate column or columns, on the same page as the relevant text (see Figure 1). A figure or table may extend across both columns to a maximum width of two columns, or 17.78 cm (7 in.).

Captions should be Times New Roman 9-point bold (`Caption` style). They should be numbered (e.g., "Table 1" or "Figure 2"), centered, and placed beneath the figure or table. The words "Figure" and "Table" should be spelled out (e.g., "Figure" rather than "Fig.") wherever they occur.

All figures should also include alt text for improved accessibility. In Word, right click the figure, and select Format Picture | Layout | Alt Text). Papers and notes may use color figures, which are included in the page limit; the figures must be usable when printed in black-and-white in the proceedings.

The paper may be accompanied by a short video figure up to five minutes in length. However, the paper should stand on its own without the video figure, as the video may not be available to everyone who reads the paper.

Inserting Images
Occasionally MS Word generates larger-than-necessary PDF files when images inserted into the document are manipulated in MS Word. To minimize this problem, use an image editing tool to resize the image at the appropriate printing resolution (usually 300 dpi), and then insert the image into Word using Insert | Picture | From File.
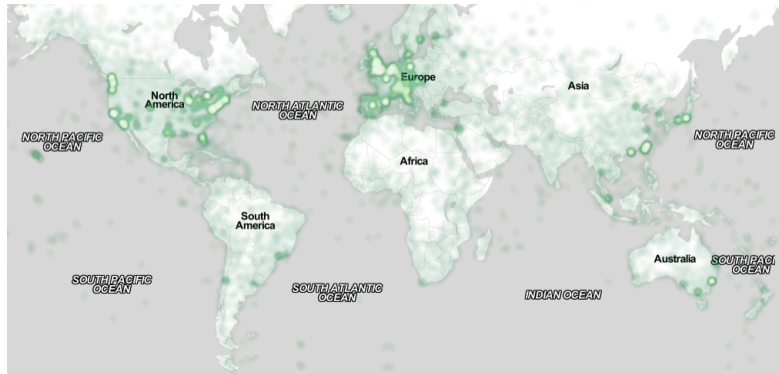

Figure 2. Sample of a wide figure. Be sure to place at the top or bottom of the page. Ensure that important information is legible in both black-and-white and color printing. Image: CC-BY-ND ayman on Flickr.

Table Style
    The text of tables will format better if you use the `Table Text` style (as in Table 1). If you do not use this style, then you may want to adjust the vertical spacing of the text in the tables.
    To adjust the spacing of text in a table in Word, use Home | Paragraph | Indents and Spacing. Generally, text in each field of a table will look better if it has equal amounts of spacing above and below it, as in Table 1. Table captions should be placed below the table. We recommend table lines be 1 point, 25% black. Minimize use of unnecessary table lines.

For improved accessibility, header rows of tables should be marked. In Word, right-click a header row, and select Table Properties | Row | Repeat as header…

LANGUAGE, STYLE AND CONTENT
The written and spoken language of SIGCHI is English. Spelling and punctuation may use any dialect of English (e.g., British, Canadian, US, etc.) provided this is done consistently. Hyphenation is optional. To ensure suitability for an international audience, please:

- Write in a straightforward style.

- Try to avoid long or complex sentence structures.

- Use common and basic vocabulary (e.g., use the word "unusual" rather than the word "arcane").

- Briefly define or explain all technical terms that may be unfamiliar to readers.

- Explain all acronyms the first time they are used in your text—e.g., "Digital Signal Processing (DSP)".

- Explain local references (e.g., not everyone knows all city names in a particular country).

- Explain "insider" comments. Ensure that your whole audience understands any reference whose meaning you do not describe (e.g., do not assume that everyone has used an Android phone, or a particular application).

- Explain colloquial language and puns. Understanding phrases like "red herring" may require a local knowledge of English. Humor and irony are difficult to translate.

- Use unambiguous forms for culturally localized concepts, such as times, dates, currencies, and numbers (e.g., "1-5-97" or "5/1/97" may mean 5 January or 1 May, and "seven o'clock" may mean 7:00 am or 19:00). For currencies, indicate equivalences: "Participants were paid ₩22, or roughly US$29."

- Be careful with the use of gender-specific pronouns (*he*, *she*) and other gendered words (*chairman*, *manpower*, *man-months*). Use inclusive language that is gender-neutral (e.g., *she or he*, *they*, *s/he*, *chair*, *staff*, *staff-hours*, *person-years*). See the *Guidelines for Bias-Free Writing* for further advice and examples regarding gender and other personal attributes [9]. Be particularly aware of considerations around writing about people with disabilities.

- If possible, use the full (extended) alphabetic character set for names of persons, institutions, and places (e.g., Grønbæk, Lafreniére, Sánchez, Nguyễn, Universität, Weißenbach, Züllighoven, Århus, etc.). These characters

are already included in most versions and variants of Times, Helvetica, and Arial fonts.

ACCESSIBILITY

The Executive Council of SIGCHI has committed to making SIGCHI conferences more inclusive for researchers, practitioners, and educators with disabilities. As a part of this goal, the all authors are asked to work on improving the accessibility of their submissions. Specifically, we encourage authors to carry out the following five steps:

1. Add alternative text to all figures
2. Mark table headings
3. Generate a tagged PDF
4. Verify the default language
5. Set the tab order to "Use Document Structure"

For more information and links to instructions and resources, please see: http://chi2016.acm.org/accessibility.

PAGE NUMBERING, HEADERS, AND FOOTERS

Your final submission should not contain footer or header information at the top or bottom of each page. Specifically, your final submission should not include page numbers. Initial submissions may include page numbers, but these *must* be removed for camera-ready. Page numbers will be added to the PDF when the proceedings are assembled.

PRODUCING AND TESTING PDF FILES

We recommend that you produce a PDF version of your submission well before the final deadline. Your PDF file must be ACM DL Compliant. The requirements for an ACM Compliant PDF are available at:

http://sheridanprinting.com/typedept/ACM-distilling-settings.htm

When creating your PDF from Word, ensure that you generate a tagged PDF from improved accessibility. This can be done by using the Adobe PDF add-in, also called PDFMaker. Select Acrobat | Preferences from the ribbon and ensure that "Enable Accessibility and Reflow with tagged Adobe PDF" is selected. You can then generate a tagged PDF by selecting "Create PDF" from the Acrobat ribbon. Test your PDF file by viewing or printing it with the same software the publisher will use, Adobe Acrobat Reader Version 10, which is widely available at no cost. Note that most reviewers will use a North American/European version of Acrobat Reader, so please check your PDF accordingly.

CONCLUSION

It is important that you write for the SIGCHI audience. Please read previous years' proceedings to understand the writing style and conventions that successful authors have used. State clearly what you have done, not merely what you plan to do, and explain how your work is different from previously published work, i.e., *the unique contribution that your work makes to the field*. Please consider what the reader will learn from your submission, and how they will find your work useful. If you write with these questions in mind, your work is more likely to be successful, both in being accepted into the conference, and in influencing the work of our field.

ACKNOWLEDGMENTS

REFERENCES FORMAT

References must be the same font size as other body text. References should be in alphabetical order by last name of first author. Example reference formatting for individual journal articles [3], articles in conference proceedings [7], books [9], theses [10], book chapters [11], an entire journal issue [6], websites [1,4], tweets [1], patents [5], and online videos [8] is given here. This formatting is a slightly edited version of the format automatically generated by the ACM Digital Library (http://dl.acm.org) as "ACM Ref". More details of reference formatting are available at:

http://www.acm.org/publications/submissions/latex_style

Note that the Hyperlink style used throughout this document uses blue links; however, URLs that appear in the references section may appear in black.

REFERENCES

1. @_CHINOSAUR. 2014. VENUE IS TOO COLD. #BINGO #CHI2016. Tweet. (1 May, 2014). Retrieved February 2, 2014 from https://twitter.com/_CHINOSAUR/status/461864317415989248

2. ACM. How to Classify Works Using ACM's Computing Classification System. 2014. Retrieved August 22, 2014 from http://www.acm.org/class/how_to_use.html

3. Ronald E. Anderson. 1992. Social impacts of computing: Codes of professional ethics. *Soc Sci Comput Rev* 10, 2: 453-469.

4. Anna Cavender, Shari Trewin, Vicki Hanson. 2014. Accessible Writing Guide. Retrieved August 22, 2014 from http://www.sigaccess.org/welcome-to-sigaccess/resources/accessible-writing-guide/

5. Morton L. Heilig. 1962. Sensorama Simulator, U.S. Patent 3,050,870, Filed January 10, 1961, issued August 28, 1962.

6. Jofish Kaye and Paul Dourish. 2014. Special issue on science fiction and ubiquitous computing. *Personal Ubiquitous Comput*. 18, 4 (April 2014), 765-766. http://dx.doi.org/10.1007/s00779-014-0773-4

7. Scott R. Klemmer, Michael Thomsen, Ethan Phelps-Goodman, Robert Lee, and James A. Landay. 2002. Where do web sites come from?: capturing and

interacting with design history. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '02), 1-8. http://doi.acm.org/10.1145/503376.503378

8. Psy. 2012. Gangnam Style. Video. (15 July 2012.). Retrieved August 22, 2014 from https://www.youtube.com/watch?v=9bZkp7q19f0

9. Marilyn Schwartz. 1995. *Guidelines for Bias-Free Writing.* Indiana University Press.

10. Ivan E. Sutherland. 1963. *Sketchpad, a Man-Machine Graphical Communication System*. Ph.D Dissertation. Massachusetts Institute of Technology, Cambridge, MA.

11. Langdon Winner. 1999. Do artifacts have politics? In *The Social Shaping of Technology* (2nd. ed.), Donald MacKenzie and Judy Wajcman (eds.). Open University Press, Buckingham, UK, 28-40.