Intro

通信的发展: 电报（Text）、电话（voice）、视频（video）、沉浸是趋势。沉浸式目前的重点，增加了更多信息。

研究delay的意义：Delay的作用，工程上很大的问题。对体验也是至关重要。所以是矛盾的关键。

Delay在2d中研究很多，还没有人研究delay的3D中的影响。

我们是第一份工作来研究这个问题的。我们怎么研究是合理的？

In this paper, we systematically reviewed prior works on Delay perception in 2D. 给了我们的启发，为了研究这个，我们build另一个具有理想延时（50ms）的系统，并开展实验了研究上述因素的影响。的感知是基于线索的，我们按照review工作，提出了一个框架，三类，它包括3类，具有同步线索的，只基于视觉的，视听觉的。我们研究在3DTI中有无audio的影响，以及同步产生的影响。

我们的贡献：

1. 系统调研和理论分析

2. 通过两个例子来验证我们的框架

3. 一套开源的系统

4. 针对同步任务的机制

5. design implication

大纲

\subsection {How users perceive network delay?}

The users perceive network delay by cues

two cases: with cue that synchronization is obvious (synchronized); with cues that can only be predicted by users with some expectation (visual, audio);

For each:

perception of local delay (how users perceive delay in synchronous tasks)

turn-talking model model (how users perceive delay through conversation)

situation awareness theory (how users perceive delay through visual feedback)

conclusion:

local delay has the highest delay requirement

turn-talking has a higher demand than situation awareness

\subsection {What is new in 3D?} (For H2)

The co-present effect.

\subsubsection {Brainstorming: Examples for each level}

Conclusion:

More visual information; Visual information matters more.

Thus, more Synchronous tasks and video-only tasks become possible in 3D.

\subsection {The less sensitiveness and more tolerance with visual information} (For H3)

grounding theory (how visual information reduce users’ dependence on conversation, for audiovisual); Previous works reveal the effect above in 2D compared to audio-only tasks; The reason can be found in grounding theory and situation awareness.

Hypothesis (conclusion):

H1: The factors that causes this effect is enhanced in 3D,

\subsection {Three levels: why we have these three levels? Can be even stronger?}

3DTI tasks can be classified into three levels according to the network delay requirement: synchronous tasks, turn-based audiovisual tasks and turn-based visual-only tasks.

synchronous tasks: definition only, no details.

turn-based audiovisual tasks: ...

turn-based visual-only tasks: ...

\subsection{Goal of the experiment}

For the first and the third, there is no related work in 2D, so we have to test.

Q1: what is the delay requirement for synchronous task?

Q2: what is the delay requirement for visual only task?

Q3: what is the delay requirement for audio-visual only task?

Q4: Is H1 supported?

\subsection {Suggestions of network design}

这下面是哪些内容？

大纲

The foremost hypothesis.

H1: 3DTI tasks can be classified into three levels according to the network delay requirement: synchronous tasks, turn-based audiovisual tasks and turn-based visual-only tasks.

synchronous tasks: definition only, no details.

turn-based audiovisual tasks: ...

turn-based visual-only tasks: ...

H2: Synchronous tasks is an emerging level in 3D, which requires a very low network delay of 50 ms. 50 ms is very challenging in a network for 3DTI. Practitioners should pay attention to judge if their application belongs to this level.

H3: Turn-based audiovisual tasks is common in 2D, which requires a network delay within 150 ms to 200 ms. However, the network delay requirement for a good user experience becomes looser in 3D. We suggest a network delay of 250 ms to 300 ms is enough.

\subsection {How users perceive network delay?} (For H1)

The users perceive network delay by cues (synchronous/turn-based \* audio/audiovisual/visual)

perception of local delay (how users perceive delay in synchronous tasks)

turn-talking model model (how users perceive delay through conversation)

grounding theory (how visual information reduce users’ dependence on conversation, for audiovisual)

situation awareness theory (how users perceive delay through visual feedback)

\subsection {What is new in 3D?} (For H2)

The co-present effect.

More visual information.

Thus, more Synchronous tasks and video-only tasks become possible in 3D.

\subsection {The less sensitiveness and more tolerance in 3D turn-based audiovisual tasks} (For H3)

Previous works reveal the effect above in 2D compared to audio-only tasks

The reason can be found in grounding theory and situation awareness.

The factors that causes this effect is enhanced in 3D, thus we propose H3.

\subsection {Examples for each level}

\subsection {Suggestions of network design}

The foremost hypothesis.

H1: 3DTI tasks can be classified into three levels according to the network delay requirement: synchronous tasks, turn-based audiovisual tasks and turn-based visual-only tasks.

synchronous tasks: definition only, no details.

turn-based audiovisual tasks: ...

turn-based visual-only tasks: ...

H2: Synchronous tasks is an emerging level in 3D, which requires a very low network delay of 50 ms. 50 ms is very challenging in a network for 3DTI. Practitioners should pay attention to judge if their application belongs to this level.

H3: Turn-based audiovisual tasks is common in 2D, which requires a network delay within 150 ms to 200 ms. However, the network delay requirement for a good user experience becomes looser in 3D. We suggest a network delay of 250 ms to 300 ms is enough.