

1.	Augmented Reality (a) What's Augmented Reality (AR) and Mixed Reality (MR)? (b) Design an AR application. Please describe in detail about the motivation, hardware/software system, and operating scenario of your design.															
	(a) Augmented reality (AR) is a direct or indirect live view of a physical, real-world environment whose elements are "augmented" by computer-generated perceptual information, ideally across multiple sensory modalities, including visual, auditory, haptic, somatosensory, and olfactory. Mixed reality (MR), sometimes referred to as hybrid reality, is the merging of real and virtual worlds to produce new environments and visualizations where physical and digital objects co-exist and interact in real time. Mixed reality takes place not only in the physical world or the virtual world, but is a mix of reality and virtual reality, encompassing both augmented reality and augmented virtuality via immersive technology. (b) The AR application I designed is an AR glasses. When the user wears the device, it will display the estimated value of the items he sees in the real life.															
7.	In terms of graphics acceleration, there are at least two approaches been proposed, namely (1) parallel processing or SIMD, such as in the NVIDIA GeForce4 chip set, and (2) binary machines such as the Pixel Planes chip set. (a) Can you describe the basic idea of how these two approaches can accelerate graphics and the major difference between them. (b) There is a function fast_calculate (A, B, C, x, y) which can be evaluated in parallel, where fast_calculate() can calculate $A*x + B*y + C$ for each pixel position (x, y). For a triangle defined by vertices (0, 6, 0), (0, 0, 0) and (8, 0, 0), please write a parallel procedure using fast_calculate() to draw all the pixels inside this triangle by Phong shading (vertex normal interpolation) with the shading function (surface normal) dot_product (light_direction). The surface normal at (0, 6, 0) is (1, 0, 0), at (0, 0, 0) is (0.6, 0.8, 0), and at (8, 0, 0) is (0, 1, 0), and the light is at (0, 0, 15). We view this triangle from Z axis above.															
8.	In order for objects in virtual environment to move realistically, the response time of the system must be very short. However, in practice, there is always system lag (latency). To solve the problem, many different methods have been proposed to "predict" the future movement upon the past motion path. One method is the simple extrapolation: <table><tr><td>Time(ms)</td><td>t - 200</td><td>t - 100</td><td>t</td><td>t + 100</td></tr><tr><td>Input</td><td>X(n - 2)</td><td>X(n - 1)</td><td>X(n)</td><td>X(n + 1)</td></tr><tr><td>Value</td><td>1.2</td><td>1.3</td><td>1.5</td><td></td></tr></table> (a) By first order linear interpolation, $X(n + 1) = ?$ (b) By second order extrapolation, $X(n + 1) = ?$ How to formulate your equation? (c) Suppose the moving object has a mass M, will this fact help the prediction? (d) Do you think more data in the past will help? Why? Is there an upper bound (number of past data) for the past data? If so, what are the factors that will limit the past length? (Hint: second order prediction: a curve passing through X(n), X(n - 1), (n - 2)	Time(ms)	t - 200	t - 100	t	t + 100	Input	X(n - 2)	X(n - 1)	X(n)	X(n + 1)	Value	1.2	1.3	1.5	
Time(ms)	t - 200	t - 100	t	t + 100												
Input	X(n - 2)	X(n - 1)	X(n)	X(n + 1)												
Value	1.2	1.3	1.5													

2.	<p>(a) Please give an example application of Augmented Reality on smartphones.</p> <p>(b) Microsoft Kinect camera can estimate the depth information of the scene in real-time. Can you give an example of research topics which could be solved or improved by using Kinect?</p>
4.	<p>(a) Is a typical RPG (role playing game) game considered VR technology? Why and why not?</p> <p>(b) What is motion parallax? How to estimate the thickness of a brick if front of us, if we have only one good eye and the other eye is blind?</p> <p>(c) Please describe three cases where it is easy to cause "motion sickness", why?</p>
	<p>(a) Yes</p> <p>(b) ...</p> <p>(c) 1. In boat rides: While you're in boat rides, the VR-device will be continually shaking, thus it'll cause motion sickness.</p> <p>2. Driving in mountain roads: Same as above, the VR-device will be shaking.</p> <p>3. Simulator sickness (SS): It's a subset of motion sickness that is typically experienced by pilots who undergo training for extended periods of time in flight simulators. Due to the spatial limitations imposed on these simulators, perceived discrepancies between the motion of the simulator and that of the vehicle can occur and lead to simulator sickness.</p>
5.	<p>In your opinion, what are the great technology breakthroughs of VR in the past fifteen years? Please list 5 such items in relative importance, and give very short explanations why it is important.</p>
	<p>1、更擬真的互動動態</p> <p>不管是 HTC VIVE、PS VR 還是 Oculus Rift，或是現有的任何 VR 裝置，在操控方面都以各種附帶的控制器進行。而不管是 HTC VIVE 的手持控制器，還是 PS4 的 Dualshock 4、PS Move，這些畢竟都是非常間接的動作體驗。想從地上撿起一個石頭？實際上我們不是真的「撿」了，而是按下控制器的按鈕，配合視覺演出想像我們把東西撿起來。</p> <p>這一點目前已經有不少廠商正研究如體感手套、動態捕捉設備等週邊搭配的硬體，來讓我們進入虛擬世界時能更擬真的互動，而不是永遠捏著兩根塑膠棒。不過說到體感相關，怎麼能忘記曾經的體感王者 Kinect？也許在微軟終於投入 VR 領域時，與 Kinect 搭配的動態表現將一口氣超車向前。</p> <p>2、與環境的動作回饋感</p> <p>解決了動態，卻還是有下一個問題，那就是各種動作的「回饋感」。什麼叫回饋感？比如說一刀揮砍下去，砍中東西時該根據目標的硬度有不同的阻力甚至反彈、比如說伸出手摸到牆壁不該是空蕩蕩的穿過去而該真的有摸到一堵牆的手感、比如說擊發槍枝的後座力等。</p> <p>在 GDC 的一場研討中，講師 Jesse Schell 提出了將有 VR 回饋機器人出現來為使用者製造這些環境回饋感。聽起來有點天外奇想的荒謬，但也不失唯一一個可能性。此外在裝備，如手套、衣服上加入微小的動力裝置去反向製造回饋感也是一個可能的研究方向。</p> <p>比如握住一個東西時，手套指套部份利用動力造成適當的阻礙，營造真的有握住東西的感覺。透過阻力的強弱、頻率、鬆弛，甚至還可能模擬出軟硬與彈性。當然，這目前僅止於空想（或者有人正在秘不可宣的研究中），而基本的回饋感如果能夠透過某個技術突破，將能讓 VR 的實感向上非常大的幾個次元。</p>

	<p>3、被侷限的移動空間</p> <p>目前已接近問世的 VR 裝置中，在空間偵測上最為理想的是 HTC VIVE，能以兩個偵測器劃出被稱為「Room-scale」的動態空間，使用者能在這個方塊空間中步行走動、轉身、蹲下起立，精確度與可動作的廣度都是其他 VR 裝置難以企及。但，終究還是在這個空間中。</p> <p>也就是說，我們光是想達到「在一望無際的花海中散步」這件事，精確來說是有難度的，我們可以透過 VR 看到一望無際的花海，但你只能往前再走三步…好吧，小步一點走個五步。如果不糾結以雙腳步行，那折衷的方式有非常多，但既然要追求極致的實境體驗，這個問題能解決會更好，是吧？</p> <p>目前讓這個問題得到解法最理想的週邊，是曾經搭配環形螢幕來進行體感遊戲的萬向跑步機 Virtuix Omni。透過專用鞋在特殊設計的底面上跑動，達成能夠朝向四面八方卻是原地跑動的動態體感。</p> <p>但 Omni 現在也有著幾個明顯的缺點，我們單純從體驗上來看撇開裝置巨大與裝備繁複的問題，首先在 Omni 機器中的使用者無法「蹲下」。因為原理上而言，Omni 是依照使用者腰部靠向某一邊來執行指令輸入，這個腰部同時還是支架功能，如果沒有想在特殊摩擦力的底盤上站定都有難度。</p> <p>但 Virtuix Omni 的作法無疑是個很大的提示，也許只要出現一個無須支架就能安全使用的萬向跑步底座，就能完全解決這個空間難題。</p> <p>4、嗅覺，難以重現的第四感</p> <p>前三點相信大家都能感受得到其必要性，以及總有一天能實現的現實感。但嗅覺？我們真的有必要連嗅覺都重現嗎？就算有必要，又真得能重現嗎？實在的說，是不是真的必要，筆者也給不出答案，但如果以追求實感的極致切入，當視覺、聽覺、觸覺都被滿足後，讓環境更擬真的下一步，毫無疑問就是嗅覺。</p> <p>戰場的血腥與硝煙味、海風的鹹味與腥味、廟宇中瀰漫的香燭味，氣味是我們構築場景印象中一個非常重要的環節，而且目前顯然是透過數據線傳不過來的東西。多年來持續聽聞有人研發出重現味道的科技裝置，但至今仍沒有出現實用化的產品也是明擺著的事實。</p> <p>而這裡其實有個有趣的延伸：如果某天嗅覺真的能虛擬實現，我們會繼續追求虛擬呈現五感的最後一項「味覺」嗎？到頭來或許《駭客任務》中那樣直接給予腦部情報的方式，在重現虛擬體驗上才最為完整又乾脆啊。</p> <p>當然，不用當電池更好。</p> <p>5. 穿戴式設備解析度的提升</p>
7.	<p>From one camera, you are looking at a triangle with known vertex positions, such as A(10, 30, 10), B(80, 30, 10), C(30, 10, 10).</p> <p>(a) How to determine the distance and orientation of the camera to the triangle?</p> <p>(b) In actual use, there are cases when one vertex of the triangle is blocked from the camera's view. What would be possible solutions to this problem? Please state your solution clearly and with explanation.</p>

2.	<p>(a) Please design a method to achieve the goal "smell real" in a movie theater? Give at least one example.</p> <p>(b) In addition to "smell real", there is "feel real". In a fiction, there is a case showing that force</p>
----	--

	feedback can be dangerous to human participants, even to the degree of being "murdered". How can this be done to a participant (user) in VR applications? Please explain.
	<p>(a) Smell-O-Vision was a system that released odor during the projection of a film so that the viewer could "smell" what was happening in the movie. The technique was created by Hans Laube and made its only appearance in the 1960 film Scent of Mystery, produced by Mike Todd, Jr., son of film producer Mike Todd. The process injected 30 odors into a movie theater's seats when triggered by the film's soundtrack.</p> <p>(b)</p>
4.	<p>Rasterization is the popular rendering technique for producing real-time 3D computer-graphics. The algorithm takes 3D scene (described as polygons) to 2D surface (usually a computer monitor). At a very basic level, rasterizers simply take a stream of vertices, transform them into corresponding 2D points on the viewer's monitor and fill in the transformed 2D triangles as appropriate. If the 3D scene is very complicated, then the number of vertices may be very large. It will take a long time to do rasterization. Consider two ways for acceleration:</p> <p>(a) The first method is to dynamically decrease the number of triangles passing to the rasterizer. That is, reduced a model from 10 million triangles to 10000 triangles. Can you propose one idea to achieve this goal while maintain the rendering quality from our view points?</p> <p>Consider the hardware acceleration using Pixel Plane, knowing that it is capable of doing the calculation <math>Ax + By + C</math> in parallel:</p> <p>(b) For a triangle defined by vertices (0, 0), (2, 4) and (6, 2) with their colors on blue channel are 10, 50, 20, respectively. Please provide the precise A, B, C values such that the result of the formula <math>Ax + By + C</math> indicates the value of the pixel(x,y) on the blue channel using Gouraud shading (smooth shading, or called color interpolation shading).</p> <p>(c) Please calculate the values of the pixels (1, 1) and (3, 2) by adopting your formula in (c).</p>
	<p>(a) ...</p> <p>(b) <math>(x, y) = (0, 0) + s(2 - 0, 4 - 0) + t(6 - 0, 2 - 0)</math>  <math>(x, y) = s(2, 4) + t(6, 2)</math>  <math>x = 2s + 6t \rightarrow 2x = 4s + 12t</math>  <math>y = 4s + 2t \rightarrow 3y = 12s + 6t</math>  <math>x - 3y = -10s \rightarrow s = -1 / 10 * (x - 3y)</math>  <math>2x - y = 10t \rightarrow t = 1 / 10 * (2x - y)</math>  Color of pixel <math>(x, y) = 10 + -1 / 10 * (x - 3y) * (50 - 10) + 1 / 10 * (2x - y) * (20 - 10)</math>  Color of pixel <math>(x, y) = 10 - 4x + 12y + 2x - y</math>  Color of pixel <math>(x, y) = -2x + 11y + 10</math>  <math>A = -x, B = 11, C = 10</math></p> <p>(c) Value of (1, 1) = <math>-2 * 1 + 11 * 1 + 10 = 19</math>  Value of (3, 2) = <math>-2 * 3 + 11 * 2 + 10 = 26</math></p>
5.	<p>3D sound</p> <p>The goal of 3D sound is to simulate the direction and distance of a sound source in a real world.</p> <p>(a) The DVD standard uses the Dolby AC-3 as its sound output, where AC-3 uses 5 speakers plus one bass speaker, as shown in Figure (a). However, your PC has two speakers only. How would you</p>

	<p>design your PC to play a DVD movie with a sound similar to Dolby AC-3?</p> <p>(b) Noise cancellation application. The newest version of BOSE headphones has the feature that can reduce the engine noise while in the jet airplane. Can noise cancellation be done? How?</p>
6.	<p><b>Human Factors</b></p> <p>If you are required to design an IMAX movie theater which adopts an LED screen instead of a project screen. The screen size is 11m high and 20m wide. Given that the field of view of human eye is 90 degrees in the horizontal direction and 60 degrees in the vertical direction, can you define the appropriate distance between the first row seats and the screen? Please calculate the least number of LEDs that should be used in this screen.</p>
7.	<p><b>Force Display</b></p> <p>(a) Please design a system with a force feedback to simulate a case in stirring an oil tank with viscosity of 30 Newton-sec/meter, while using a Microsoft Future Force Feedback joystick with a mass of 0.3 Kg, stiffness of 0.01 Newton/meter, and viscosity of 1.17 Newton-sec/meter. Please write the necessary equations for the necessary force to be generated from the driving motors inside the joystick.</p> <p>(b) There exists one magic trick about the above system. When simulating a spring with stiffness of 1000 NT/m, you are holding the handle of a force feedback joystick. The harder you tense your forearm in holding the handle of the joystick, the stronger is the vibration in your hand. Can you explain this magic behavior? At what sampling frequency can this happen? (Hint: Human hand has stiffness of 400 NT/m, viscosity 5 N-sec/m, and mass 2Kg, while the joystick has a viscosity of 1.5 N-sec/m).</p>

3.	<p>In terms of graphics acceleration, there are at least two approaches been proposed, namely</p> <p>(1) parallel processing or SIMD, such as in the NVIDIA GeForce4 chip set, and</p> <p>(2) binary machines such as the PixelPlanes chip set.</p> <p>(a) There is a function <code>fast_calculate(A, B, C, x, y)</code> which can be evaluated in parallel, where <code>fast_calculate()</code> can calculate <math>A*x + B*y + C</math> for each pixel position <math>(x, y)</math> in PixelPlanes. For a triangle defined by vertices (5, 9), (3, 2) and (10, 3), please write a parallel procedure using <code>fast_calculate()</code> to draw all the pixels <math>(x1, y1)</math> inside this triangle. (Hint: you have to determine whether <math>(x1, y1)</math> is inside of the triangle first, and for this case, what are the value of A, B, and C for this particular triangle. In general, what is the distance from <math>(x1, y1)</math> to the triangle edge?)</p> <p>(b) For 3D triangle vertices A(5, 9, 0), B(3, 2, 0) and C(10, 3, 0), how to use Z-buffer algorithm in Pixel Planes for visibility testing? Hint: Z-buffer algorithm:</p> <p>(a) Initialize <code>frame_buffer[1000, 1000]</code> to be all very big numbers (far from view point). For a pixel <math>(x, y)</math> located inside a triangle,</p> <p>(b) If it is closer (its Z value) to the view point (human eye), then paint this pixel with its color. At the same time, set <code>frame_buffer(x, y) = Z value</code>.</p> <p>(c) Otherwise, don't paint this pixel.</p>
----	--

4.	<p>Assume that a Google glass VERSION G has</p> <ul style="list-style-type: none"> <li>(i) camera</li> <li>(ii) lens on the right eye, that can reflect images from a Android phone, through a prism, for head-up display</li> <li>(iii) a Microsoft Kinect depth camera.</li> <li>(iv) Blue-tooth for near field communication</li> <li>(v) a GPS sensor.</li> </ul> <p>Hint: how to make the images from Android phone to be projected to your right eye through a prism?</p> <ul style="list-style-type: none"> <li>(a) Although the exact configuration is not yet disclosed, please draw a design (sketch diagram) so that every module is connected, and in its correct positions.</li> <li>(b) Please give one example of its possible cool usage, especially those that are not yet mentioned in the Internet introduction.</li> </ul>

102

1.	<p>The da Vinci System is designed to facilitate complex surgery using a minimally invasive approach. It consists of a surgeon's console that is typically in the same room as the patient, and a patient-side cart with four interactive robotic arms controlled from the console. Please answer the following questions:</p> <ul style="list-style-type: none"> <li>(a) What technologies related to Virtual Reality are involved in this system? Please list 3 items and give a brief explanation.</li> <li>(b) How can the surgeon know that the robotic arms had touched the patient? What good things will you design to inform the surgeon about the above situation?</li> </ul>
3.	<p>In terms of graphics acceleration, there are at least two approaches been proposed, namely (1) parallel processing or SIMD, such as in the NVIDIA GeForce chip set, and (2) binary machines such as the PixelPlanes chip set.</p> <ul style="list-style-type: none"> <li>(a) There is a function <code>fast_calculate (A, B, C, x, y)</code> which can be evaluated in parallel, where <code>fast_calculate()</code> can calculate <math>Ax + By + C</math> for each pixel position <math>(x, y)</math> in PixelPlanes. For a triangle defined by vertices <math>(1, 2)</math>, <math>(4, 4)</math> and <math>(7, 0)</math>, please write a parallel procedure using <code>fast_calculate()</code> to draw all the pixels inside this triangle.</li> <li>(b) For the triangle defined by vertices <math>(1, 2)</math>, <math>(4, 4)</math> and <math>(7, 0)</math> with their colors on blue channel are 17, 31, 21, respectively. Please provide the precise A, B, C values such that the result of the formula <math>Ax + By + C</math> indicates the value of the pixel <math>(x, y)</math> on the blue channel using Gouraud shading (smooth shading, or called color interpolation shading).</li> </ul>
4.	<p><b>Motion Capture</b></p> <p>Motion capture is the process of recording the movement of objects or people. Traditionally, the captured subject would wear several markers and the related motion would be recorded by several surrounding cameras. Nowadays, Kinect uses depth cameras to track human skeleton motion.</p>



	<p>Hint: Kinect's depth acquisition is enabled by "light coding" technology. The process codes the scene with near-IR light, light that returns distorted pattern depending upon where things are. The solution then uses a standard off-the-shelf CMOS image sensor to read the coded light back from the scene using various algorithms to triangulate and extract the 3D data.</p> <p>(a) Compare the traditional maker method and Kinect system. What are the pros and cons of each method?</p> <p>(b) Assume that you can capture human skeleton motion successfully. Please give one example of the possible cool usage, especially those that are not yet mentioned in the Internet introduction.</p>
6.	<p>In Windows 8, a multi-touch screen similar to Apple's iPhone is one unique way to provide new user interfaces. Assume that we have "two" cameras in each Notebook PC, located at the top right and left corner of the screen.</p> <p>(a) Can you describe a way to provide "multi-touch" user interfaces without even "touching" the screen?</p> <p>(b) Please describe your algorithm to solve the problem of determining finger positions (not just one, but perhaps two fingers).</p>

103

2.	<p>Head-Mounted Display</p> <p>Assume you are an engineer sitting in the office and you are anxious about the Google I/O in 2015 because you are in charge of designing Google Cardboard Part 2.</p> <p>(a) You start from hardware of Google Cardboard. Explain why you need two lens for your device. (Hint: What kind of lens will you need? What are the functions of the lens? )</p> <p>(b) After assembling all components, you start to design the software. You put figure 1 in your cellphone and start to watch it, but you found it to be blurry. What kind of adjustments for image will you need in your software? (Hint: What is the effect by lens and binocular vision?) (Figure 1: Same image in both left and right side)</p> <p>(c) You can see an image without blur but it is not with enough resolution. The length and width of your screen is 12 and 6 cm respectively and the distance between your eyes and screen is 6 cm. How many points will you need in your screen at least to achieve the best resolution? (Hint: Visual acuity: eye's ability to distinguish two points of light is limited to 2.0 mm at a distance of 10 meters.)</p> <p>(d) Finally, you start to test your device. By using the gyroscope in the cellphone, you can pitch, roll and yaw. If you want to increase another inputs besides the gyroscope, please describe how to make the cardboard device obtain such information? Describe your methods in brief.</p> <p>(e) If you are about to publish Google Cardboard Part 2 next week. What kind of applications will you use?</p>
3.	<p>Augmented Reality and Virtual Reality</p> <p>(Figure 2: Oculus Rift(Left) and Microsoft HoloLens (Right))</p>

	<p>VR and AR are popular in recent years. Oculus Rift and Microsoft HoloLens are two such products which have high publicity.</p> <p>(a) What's the difference between Augmented Reality (AR) and Virtual Reality (VR)? Please classify the products (Oculus Rift and MSFT HoloLens) and explain it with your reasons.</p> <p>(b) One of core functions in Oculus is how to accurately determine the user's head movement. In order to improve user experience, Oculus uses a camera to assist it. Please explain why this camera can help Oculus to determine the user's head movement. Also, describe possible problems of this camera.</p> <p>(c) In HoloLens promotional video, there is a scenario that a woman is guiding by her father how to install the pipeline with some helpful virtual information such as installation path (Figure 3). Please analyze what kind of technology will be used in order to achieve the above result.</p>
6.	<p>Suppose we are going to use only two speakers to implement a 5-speaker Dolby AC-3 sound system in your study room. What will be the technical problems you have to solve instead of using headphones only? Give your formula or algorithms. Tell us what's missing (difficult to calculate) in your solution?</p>

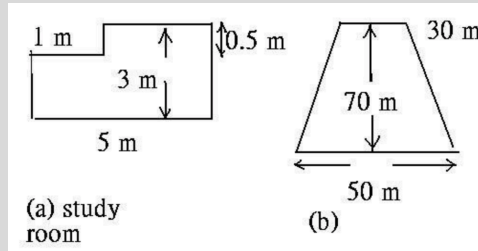
## 106

1.	<p>VR Introduction</p> <p>(a) What's the major four problems in Augmented Reality (AR) and Virtual Reality (VR)?</p> <p>(b) Given a Google cardboard HMD (head mounted display), please list its basic features and functions. How can these be achieved? Please draw pictures to explain.</p> <p>(c) Please list two AR applications not yet mentioned in our class notes.</p>
	<p>(a) 1. Display resolution: wide angle display, without seeing pixels in display  2. Latency: should be less than 20 to 50ms If greater than 50 to 100 ms, will cause dizziness, even nausea (作嘔) .  3. Fixed focus vs. variable focus in observation through HMD or others.  4. Registration of real objects and virtual environments in AR.</p> <p>(b) Independent left and right eye view.  Using your smartphone's hardware. Tracking device: orientation + position.</p> <p>(c) World Brush: 帶著這個刷具，你可以造訪世界各地名勝用它留下各種塗鴉紀念。完全合法！不會破壞任何自然景觀。  Fitness AR: 不論是晨跑、環山、健走，現在你可以透過 3D 圖記錄自己行走的路徑，Fitness AR 透過 3D 顯示出陡峭或是平緩的道路，讓你更容易了解與調整你的運動範圍。</p>
2.	<p>Motion sickness and Display Resolution</p> <p>(a) What's Motion Sickness (simulator sickness)? Which human organs are involved in motion sickness? And what are the possible factors in VR simulation that will cause simulator sickness? (14pts)</p> <p>(b) Given a new TV set (70 inches diagonally, 16:9 Aspect ratio) in your room, with a resolution of 4K (4,096 x 2160 resolution). What's the minimum distance for you to watch this TV without</p>



	seeing pixels on the screen?
	<p>(a) When you use an AR/VR device for a long time (5mins ~ 15mins), you feel some kind of dizzy. Eyes + Ears. Latency</p> <p>(b) Eye's ability to distinguish two points of light is limited to 1.5 – 2.0 mm at a distance of 10 meters. (P52) 1 inch = 25.4 mm 70 inches = 1778 mm Set width = 16x, height = 9x -&gt; <math>\sqrt{(16x)^2 + (9x)^2} = 1778 \text{ mm}</math> -&gt; width = 16x <math>\approx</math> 1550 mm Each pixel = 1550 / 4096 = 0.38 mm For 2.0mm (眼睛比较好的): <math>0.38 / 2.0 * 10 = 1.9 \text{ m}</math> For 1.5mm (眼睛比较差的): <math>0.38 / 1.5 * 10 = 2.533 \text{ m}</math></p>
3.	<p>Motion Capture</p> <p>(a) What's difference between Outside-in and Inside-out tracking (refer to reference papers)? What are the pros and cons of each method?</p> <p>(b) Assuming we have two marker's position in world space and camera's intrinsic parameters, can we find the pose (rotation and translation) of the camera? If we can, how?</p> <p>(c) Can a typical Google Cardboard VR measure the player's translation? What are the possible shortcomings of this method?</p>
	<p>(a) P167</p> <p><b>Outside-in tracking:</b> Many cameras (commercial devices) are looking at you. Pros: You can randomly put the cameras. There are markers everywhere, so you have a large walking room. Cons: You are limited to a particular space.</p> <p><b>Inside-out tracking:</b> A human being is wearing cameras looking outside. Pros: Cons: The markers are fixed. You cannot change it anywhere. LED is inexpensive &lt;-&gt;</p> <p>(b) ...</p> <p>(c) Integration over a long time, the accumulated errors can be very big.</p>
4.	<p>3D Sound</p> <p>The goal of 3D sound is to simulate the direction and distance of a sound source in a real world.</p> <p>(a) Please explain how to reduce the (big) engine noise in an airplane for the pilot wearing a headphone, while listening to music.</p> <p>(b) The following is the shape of your study room, and you want to hear the music (without wearing a headphone) exactly as it was generated in the opera house in Paris, France. Can it be done? How?</p>

Hint: use head related transfer functions (HRTF) and its measurements, then write down each steps to achieve this goal.



(a) We can measure the noise to a sine wave, then reversing the phase of the wave to get a negative wave and finally we can cancel the noise.

(b)

Pans opera house: A

Your study room: B

If we can send people to measure A, then

$$X_1 * L * F_L * M_L \text{ (left ear)}$$

$$X_1 * L * F_R * M_R \text{ (right ear)}$$

where  $M_L = M_R$  (Microphone transfer function), L: loud speaker.

Similarly, in your study room, measure

$$X_2 * L_S * F_{SL} * M_S \text{ (left ear)}$$

$$X_2 * L_S * F_{SR} * M_S \text{ (right ear)}$$

Where  $M_S$ : Microphone,  $L_S$ : loudspeaker.

Therefore, for the right ear signal:

$$X_2 = (X_1 * L * F_R * M_R) / (L_S * F_{SR} * M_S)$$

## 6. Force-feedback devices

Allen build his 3D sculpting system with a force-feedback device.

The pseudo code of his main algorithm is following:

```
while(sculpting) {
    Do the "collusion detection" with current sculptor position
    Calculate the "feedback force" and set it to force-feedback device
    Modify the 3D scene
    Render the 3D scene
}
```

Assuming the joystick stiffness = 0, mass = 0.5 KG, and viscosity = 4 NT-Sec/Meter.

While human operator (arm and wrist) stiffness = 400 NT/m, mass = 3 kg and viscosity = 3 NT-Sec/Meter.

(a) Let's simulate a spring (with spring constant  $K = 200$ ) with a length of 40 cm, while at both ends are "rock" walls (hard surface) using such a spring system using a large spring constant  $K = 500000$  NT/m. Your joystick is at the middle of the spring, say at 20 cm. Please write down your simulation equations.

(b) Do you think it is a good system design? What are the potential problems? Hint: (for a system to be stable: Sampling period  $T \leq 2 * B / K$ , where B is the total system viscosity, K is the total system stiffness.)

	(c) For the problem in (b), can you think of ways to actually simulate a hard surface in virtual reality, and still feels “realistically”?
	<p>(a) <math>Ma + bv + kx = F_s - F_{\text{ext}}</math>  <math>F_s = (0.5 + 3)a + (4 + 3)v + (0 + 400 + 200 * 4) + F_{\text{ext}}</math> (40cm: 200, half: *2, parallel: *2)  (P212 – P214)</p> <p>(b) The system becomes unstable, when <math>KT - 2 * B &gt; 0</math>, where K = stiffness of the system, B is viscosity, and T is the delay.  Actually <math>T &gt; 2 * B / K</math> can be unstable for the system! (P219)</p> <p>(c) (P220)</p>
7.	<p>Term project</p> <p>What is the term project you want to do for this semester? What kind of technical problems will you encounter and have to be solved?</p>