

A discussion on the technology used in Music of 9

Abstract

The following is a reproduction of the discussions that occurred during a meeting that was held remotely on 5 May 2022 during the presentation period of Music for 9, an exhibition under The Missing Link project.

In Music for 9, video of a percussion performance is fragmented, visually and aurally. Close-ups of body parts will be shown on synchronised displays, while the speakers that are paired with the displays will fill the space with impure sounds. The entire visual and audio spectrums are filled with rhythms and melodies built by re-integrated moving images and sound. Sound will gradually change from solely percussion music to sounds of the city. While the images on the screens remain unchanged, the sound that gradually breaks away from the image will take you to imaginations of the city soundscape.

The following discussions were translated from the original Cantonese into English and edited for clarity. The participants were Wesker Shek (software design and installation), Fung Wing-lam, Cat (video and audio synchronisation design), Vanissa Law, (co-curator, the Missing Link), Yan Lam (project Manager, the Missing Link), and Wilson Chau (project assistant, the Missing Link).

The Robot Dog

Vanissa: Let's start with the robot dog. The robot dog used (Fig. 1) is a Go1 EDU manufactured by Unitree Robotics. This is designed to be a consumer product to amaze us with its dog-like behaviour. On the Unitree website, it claims to be able to accompany you while you are jogging or walking in the park, or even bring you your morning paper. Apart from its consumer features and affordable price, Unitree also emphasises its programming and development opportunities. How programmable is the Go 1 EDU, Wesker?

Wesker: This robot dog is a product still under development. There are two ways to access the consumer features of Go1 EDU, either through the Unitree Robotics app that runs on Android and IOS

platforms, or through the tag controller (which looks like an ordinary game controller with a joystick) that comes with the robot dog. However, we used neither of the above means because we want to control the robot dog through the Robot Operating System (ROS).



Figure 1. Robot dog Go1 EDU manufactured by Unitree Robotics.

There are some existing libraries available online for people to access the control of the dog without going through the official ways. However, some of the features, such as the built-in torch, camera, and motions sensors, are locked. Therefore, we have no way to know about the surroundings. The only controllable part is the motor group, which means I can only control the dog to walk, sit, or turn (mobility). Because the built-in features are not available, we will have built our own external sensor devices. Currently the programme for these is run on a Raspberry Pi, which is mounted on the back of the robot

dog by a custom designed 3D printed part. A compass is used to locate the dog's orientation.

Object detection by colour tracking

Wesker: A second Raspberry Pi with an Intel RealSense D435i depth sensor is mounted on the cage, also by a 3D-printed component, to locate the position of the dog. The D435i is a stereo vision depth camera system that provides RGB and stereo depth resolution. It also includes Inertial Measurement Unit (IMU) for 6 Degrees of Freedom (6DoF) data. In the beginning we programmed an AI to do object detection of the dog through Machine Learning using You Only Look Once (YOLO). YOLO architecture is a popular algorithm for object detection developed by Joseph Redmon et al. in 2016. The version used in this work being YOLO v7, which is believed to be currently the most powerful object detection algorithm.

It was not very efficient when we tested it at the actual exhibition space, and we suspect the following reasons: firstly, the colour of the body of the dog and the floor are too similar, which make it very difficult to do edge detection. Secondly, because the exhibition space is surrounded by big floor-to-ceiling windows (Fig. 2), the changes in ambient lighting of room makes colour tracking very challenging.



Figure 2. Exhibition space surrounded by floor-to-ceiling windows.

In fact, object detection with optical cameras, even when paired with IR cameras, depends deeply on environment lighting. Therefore, it is very important to have sufficient time to test and debug at the venue. This is one of the very common challenges in the use of technology in art, that is, not having enough time to try things out at the venue.

Vanissa: I personally have experienced the same problems when trying to do colour tracking and object

detection with RGB and IR cameras. Gesture detection hardware such as Kinect are usually more reliable because the human body is warm enough to emit detectable infrared radiation. I reckon the difficulties in tracking the dog might be due to the lack of body temperature of the dog.

Wesker: I think that might be one of the factors. After discussing with GayBird we tried do the tracking by QR codes. A camera is mounted on the inside of the cage, which will locate the QR code that is attached to the back of the dog. This did not go well because of extreme changes in ambient lighting at the venue. Because of the floor-to-ceiling glass window, luminosity of the room changes with the outdoor sunlight conditions. In some light conditions, the white on the QR code was recognised as grey, and therefore led to very unstable tracking. We tried to put different lighting around the cage to keep the ambience light level steady, so that the camera can operate under a relatively stable lighting condition, but we still failed. Because we only had a week to set up, we eventually gave up position tracking completely, and instead put a leash on the dog to limit its range of movement, to make sure it would not crash into the cage..

Programming the movement of the dog

Vanissa: What control of the dog did you and GayBird want initially?

Wesker: Initially GayBird asked me to access all available functions and build a user interface for him so that he could control the dog through Open Sound Control (OSC). He changed his mind later and thought the dog should move randomly in the cage instead. Hence, the movements of the dog are randomised in the current design. There is an interesting fact about the dog though: that is, no matter how you ask it to move, it tends to favour in moving to its right.

Vanissa: So how did you hack into this robot dog? Are there any existing libraries available?

Wesker: I did that through the Robot Operating System (ROS), and there are some libraries available online, but they are quite buggy because they are still under development. There are two types of control available: low-level, and high-level. Low-level controls are those that control the operation of each individual parts and motors. But if we want the dog to perform certain gestures such as turning, kneeling, and walking, it

requires combinational control of several motors simultaneously.

Vanissa: So, the library we used now provides both low-level and high-level programming control?

Wesker: Yes, it provides both, but the user manual of the library is not thorough. We had to learn about what control we had access to by performing some tests. I took the dog to Cat's studio to perform some tests based on the manual. After a few days we could access high-level (gesture) control through the library. Not all the functions are accessible though; for example, I cannot ask the dog to sit down.

Vanissa: I see. That means there is a function to command the dog to sit in the app, but it's not accessible when you try to hack it with ROS.

Wesker: Yes. A week before the exhibition opened, GayBird asked me to try to make the dog perform random gestures through low-level (single motor) controls. While I was working on that in the exhibition space, suddenly, the dog produced a loud beeping sound that sounded like a bomb was about to go off. It was so loud that the whole venue was filled with it. It was the first time I heard the dog make sounds like that. We had no idea why it did that. GayBird then emailed the manufacturer and asked; they replied and said the low-level controls cannot be accessed through programming. It can only be done through a specific controller that is sold as an accessory of the dog. That controller looks like a regular game controller. One is supposed to access low-level control by keying in a specific series of command. We then tried it with the controller, and it performed exactly what the manufacturer described.

Vanissa: Oh, so you guys do have the controller?

Wesker: Yes, we do, but we do not want to use the controller as it would defeat the purpose of the exhibit. We wanted to map and automate the gestures of the dog so that the dog will respond to its surroundings. If this can only be done through a controller operated by some person, the whole purpose is lost. And this "series of commands" on the controller to activate low-level controls is very clumsy. We worried that the helpers, who are supposed to turn the exhibit on-and-off every day, could not manage to do this, therefore we gave up on using the low-level controls.

We now only focus on using high-level controls to design the response of the dog. The dog will now kneel, stretch its limbs, and other gestures like an actual dog will do.

Then GayBird said those gestures are what he was looking for, therefore we settled with this current design.

Vanissa: It is always challenging to hack some commercial products, especially the product is still under development. Any other thing you want to talk about, Wesker?

Power supply of the dog

Wesker: I would also like to talk about the problem we had with the power supply of the dog. The original plan was to put the dog on an adaptor so that it will operate full-time during exhibition hours. According to the manual, the dog runs on 24-volt, 30-amp DC (direct current). We have done several tests with different adaptors, but the power provided is not enough to make the dog stand up, even with the adaptors that run on higher current. There were a few times it stood up for just one second and lay back down immediately. The manual says the dog will check its power supply upon booting and, if the power is not sufficient, it will stand up for one second and lay back down immediately, exactly what happened during the test. Based on this we assume it is a power supply problem.

GayBird said he saw some videos of this robot dog operating normally on a power adaptor. We suspect we might either have a different model from the ones that is used in the videos or, more likely, the internal chargeable battery inside of the dog is dead. I am not sure if it is a factory defect, or if it is damaged or aged after a few years of usage. Now it can only be run on external portable batteries, which are only sufficient to power the dog to perform for one hour each time.

Vanissa: Yes, I recall, at one point, we talked about getting three-phase alternating current (AC) from the venue to power the dog?

Wesker: Yes, we did. Three-phase power is a three-wire AC power circuit with each signal 120 electrical degrees phase apart. Residential homes use single-phase power supply, while commercial and industrial facilities usually use a three-phase supply. The key difference between the two is that a three-phase power supply is better in accommodating higher loads. We did request for three-phase power because we thought the dog might need very high power to run. The dog runs on 24 V and 30 A, which is already 720 W. And a normal one-phase power plug provides only 220 V and 10 A, which is around 2200 W. Hence, before we figured out the problem with

the internal battery of the dog, we guessed that the unpredictable performance might be due to insufficient power. Other parts of the exhibition, such as computers, display screens, interfaces, mixers, and speakers, also require a lot of power to run. We ended up not using the three-phase power supply when we figured out it was not a problem of insufficient power. We chose to run the dog on an external portable battery instead.

Synchronisation system

Vanessa: Let's move on to the synchronised audio-visual system that is built by Cat. How many channels of audio and video are there in total, Cat?

Cat: There are nine video displays at the venue. Seven are in the main exhibition space, each paired with its own speaker (Fig. 3). The seven transparent video displays in the main exhibition area, along with two LED TV in the small room, are connected to BrightSign media players, while audio in the two rooms is from different sources.



Figure 3. There are seven monitors and speakers in the main exhibition area.

Vanessa: Oh, the two spaces have audio from different sources. Are they still synchronised?

Cat: Yes, they are, and this is where the challenges are. We use BrightSign media players to synchronise the videos, and on each unit, there is a 3.5 mm audio jack output for audio. GayBird was not happy with the quality of the audio, and asked if we could find ways to synchronise audio outside of the BrightSign network, so that the maximum audio quality can be preserved.

Vanessa: Sorry to interrupt. Can we step back and talk about the BrightSign media players a bit? Is this a piece of hardware?

Cat: Yes, BrightSign is a piece of hardware that deals with one-channel video and one-channel audio. When

more than one unit is used, synchronisation between the units can be achieved through a UDP protocol. BrightSign is a leader in providing digital signage media players. They have a reputation of its ease of use, reliability, and affordability (Fig. 4). Theoretically, if we are using a total of nine display monitors at the venue, we will need nine BrightSign units. Although they are relatively affordable, around HKD 4000 per unit, compared to other video synchronisation methods, it would be quite expensive if we needed nine of them. We found ways to get around this and now we are using only three BrightSign units in the exhibition. Two are used in the small room, with the two TVs; two are used in the main exhibition area with the seven transparent video displays (Fig. 5). Because the transparent displays are small, and they do not require a very high video resolution, we split a 4K video into four 720-pixel videos through a splitter, so that only one BrightSign is required for four transparent displays. This can be achieved only because the BrightSign units delivers stellar "True 4K" video playback.



Figure 4. A BrightSign media player unit.

Synchronisation of audio and video

Vanessa: That's very smart solution! Nine units versus four units! So how is the synchronisation between BrightSign and audio done?

Cat: Audio came from an Ableton session, which is the master in this audio-visual network. Audio is sent from Ableton to the nine speakers in the venue. We use Max4Live in Ableton to send OSC messages to the three BrightSign units to trigger video playing. The BrightSign units are slaves in this network waiting for commands from Ableton.

Vanissa: So, the videos are stored in the BrightSign units?

Cat: Videos are stored in SD cards in each BrightSign units, together with the BrightAuthor:connect software. All the units, together with the computers, are connected to the same router, with their own IP addresses. Commands (OSC messages) are sent from the master to the software on each unit (slave), to control which video to play, and how the video should be played.

Vanissa: Is the computer that is running Ableton also running the BrightSign software? Are there any real-time responses? Is this because you have to plan ahead and upload all the code onto the BrightSign SD cards beforehand?



Figure 5. Transparent display with speaker in the main exhibition area.

Cat: No. The computer is not running the BrightSign software. The OSC messages are sent from Ableton to the BrightSign software to tell each unit what to do. They will work if they are all connected to the same network. And yes, programming needs to be done and uploaded to the SD cards beforehand. The programme sets up each BrightSign unit by telling them information such as file names of the videos, what messages they are going to receive, and what to do upon receiving those messages.

Using BrightSign with computer

Vanissa: I see. The control is real-time. What needs to be planned ahead in the programme is what messages to use. Once the SD cards are in the BrightSign units, you cannot change the programme unless you take them out and upload the code again.

Cat: Yes. We did change the code a few times because we need to fine tune the framing of each split videos. We use HDMI, which is a proprietary audio/video interface. For the transparent displays, because the audio is from

Ableton, we will need to mute the audio on the BrightSign. It has to be done in the code, and there were times we forgot to mute the audio and we had to re-do the code and upload again. We did de-bug and fine-tune the code several times during set-up.

This is the first time I used BrightSign in this way. I usually connect them through a master-slave network, in which one of the BrightSign units is the master. If you connect them this way, no computer is needed. The master unit will handle all the control. It is actually quite handy because once they are connected to a power source, they are good to go. It takes them a while to boot, then they will find each other, connect themselves, and start playing videos on their own. This was my first time using a computer as the master; only when GayBird requested the audio be sent from Ableton instead of the BrightSign units, I discovered connecting the units this way.

Vanissa: Is the frame-dividing also done in BrightSign?

Cat: The frame-dividing is done by the splitters. We used four BrightSign units in total, two of which are in the small room for the big screens. The other two, which are in the small room, are connected to a splitter. The splitter has one HDMI input and four HDMI outputs. It provides different settings for how the frame is split.

Vanissa: I see. What about audio in the small room; does it come out of the 3.5 mm audio output on the BrightSign units?

Cat: Yes, audio output in the small room is from the BrightSign units. GayBird uses an oscillator as a pre-amp for the audio signal that runs from the 3.5 mm audio output.

Vanissa: Does the audio that comes out from the BrightSign units need pre-amplification?

Cat: I guess a pre-amp is needed because GayBird wants to use a speaker cone; therefore, it needs a driver or a pre-amp.

Importance of the power up sequence

Vanissa: Next I would like to invite Yan and Wilson to discuss if there are any technical difficulties that you think are worth mentioning in this report.

Wilson: There were several times when the system was booted that the videos were out of sync. I think Cat was there.

Cat: Yes, I was there. This is about the messages sent to the BrightSign units. You can send messages to BrightSign to tell it to start playing, or stop playing, a specific video. The video starts to play when the units received a “start playing” message; if they receive another “start playing” message while the video is playing, the video will not be played from the beginning. To fix this, a “stop playing” message needs to be sent to the unit to start everything over again.

(software or hardware) is updated. For example, when the codes in the BrightSign units are updated, we must test that the switching-on-off sequence still works.

Vanissa: The displays on the “alter” that shows the back of the head of GayBird, are those also connected to a speaker?

Cat: Yes, two speakers are connected to that screen. Audio also came from Ableton through an interface and mixer, while audio of the speakers on the floor came through an interface and pre-amp.

Sending live video through network

Vanissa: TouchDesigner is running on the computer in the small room. What is it responsible for? OSC?

Cat: TouchDesigner is responsible for the projection in the main area. This is a collaborative work of Wesker’s and mine. GayBird wants to project the vision of the dog on the wall opposite to the “alter.” The projector is in the main area; however, the input of that projector is in the small room. We need to find ways to send live video from the webcam on the dog into the small room. If we did this with cable it would require an almost 30-meter cable. Technically there is no specific maximum length a cable can go, but it depends on the type and material; cable of this length will run into its limitation and tend to create problems. For example, copper cable loses signal strength (attenuation) the further it reaches. For cable that is longer than ten meters, usually an extender or repeater will be needed, or one can switch to fibre-optic HDMI cables that have smaller attenuation. For the webcam we therefore ditched cable and sent video through the wireless network instead. The webcam on the back of the dog is connected to the computer in the main area. Live video is sent from that computer to the computer in the small room through the network using TouchDesigner.

Vanissa: And there is some video processing that is applied to the dog-view live video feed. Is it some native function in TouchDesigner?

Cat: Yes, that’s a mask video effect in TouchDesigner. The edge of the video is blurred with an oval shape, so that it looks more like a low-angle view and gives it some dreamy feeling. In the original frame, the faces of the audience can also be seen.

Vanissa: Is there anything else worth mentioning?

開門Checklist (先開電掣, 再開電腦)	
正門入口	左右兩邊玻璃門需要開鎖 - (逢星期三至五) 請出入開門 - (逢星期六、日及公眾假期) 請全日打開左邊的門 請掛上門牌在左邊門框上
士多房	開燈掣 (藍色範圍內, 總共六個) 開冷氣掣 (藍色範圍內, 總共四個) *LP 7L2* & *LP 9L3* 兩輪掣須長開保持著啱 (紅燈) 須把寫著“MONITOR”的黑色插頭插到牆上
細房門口	長槍槍底插電掣
大房電腦	開白色拖板開關->Orange switch 兩個->Green switch 總共四個制 (請清楚明燈是否著啱)
大房audio	拖板 由右至左開(one by one no rush)
大房電腦	開laptop-Ableton Live suppose會自動開 (如果另一個app唔啱自己揀返Ableton Live唔啱)
細房電腦	開Dell-應該會見到TouchDesigner show緊機械狗視野 *Dell Winon*需要選擇 需兩個remote [Panasonic] [NEC]出去大房開projector+大電視
大房	用[Panasonic] remote(左上角)開機械狗視野projector
近正門入口	用NEC remote(左上角)開 Artist Statement 最大電視
大房電腦	in Ableton Live (1) press A (2) press play 跟住應該聽到聲, 都睇睇所有screen 有冇visual
機械狗開機	power 按一下短 一下長 (入籠小心頭小心電腦)
機械狗換電	入籠小心頭小心電腦 機械狗開機: power 按一下短 一下長 (留意機械狗都幾重, 開機要抱實) 房較左手能動位置 換電 放返隻狗喺正中交叉位(patpat 係mark 上面) 屋返啲手腳去白切雞位置訊好 (make sure 放好, 唔係隻狗會都唔啱)
關門Checklist (先關電腦, 再開電掣)	
大房電腦	in Ableton Live press stop, close all app, shut down macbook and close the lid
細房電腦	press esc, close the app, shut down dell computer
細房門口	長槍槍底unplug
大房	turn off projector (Panasonic) turn off BIG TV (NEC)
機械狗開機	power 按一下短 一下長 (記得抱實隻狗因為佢好重) turn off 4 power switches next to Macbook turn off 4 power switches on the audio side(wait until red light turn off for the last switch)
Store room	turn off air con turn off lights turn off store room lights check any trash around close the store room door
Exit	close main door

Figure 6. Notes to docent about power up sequence.

Vanissa: It’s similar to MIDI. Without a “note off” message, a note will be sustained forever.

Yan: Yes, this happened several times. Cat and I then tried to work on a correct sequence to turn the hardware and software on and off. A correct sequence of switching on the system is so important because, if the software is opened before the hardware is up and running, it might never be able to recognise that piece of software. After we have a sequence of switching on equipment (Fig. 6), the problem is solved. However, we will have to test the sequence again every time when any part of the system

Cat: We can also talk about the cables on the floor. For each transparent display, there are two power cables (for the transparent display), one audio cable, and one HDMI cable, which are covered by cable slots on the floor. Because of the current design of the cable slots, two of the HDMI cables are longer than 15 meters. As I mentioned before, we will need an extender or repeater to compensate the signal loss, but that would require an extra power cable. In the end we used directional HDMI cables. Directional HDMI cables include built-in equalizers and amplification and are designed to carry signal over longer distances. One end is labelled "source," or "1," and the other end is labelled "TV," or "2." During installation we did not pay close enough attention and reversed the direction of the cable. When the videos were not running, we tried to troubleshoot and tested if the screens are faulty, or perhaps the cables, or if we were using the splitter incorrectly. Eventually we realised that we had the directional HDMI installed in reverse. This mistake costed us a lot of time because we had to remove all the cable slots, reverse the cables, and put the slots back again.

Vanissa: This is the first time I have heard of one-way HDMI.

Cat: When the length exceeds 15 meters, this type of HDMI cable is available.

Vanissa: Oh, I have a question about that. Is it difficult to put the HDMI cables into the cable slots because HDMI cables are usually quite stiff, not easily bent?

Cat: Not really. We did not have problem putting them into the slots. HDMI cables do indeed need extra care though. In an exhibition setting, if the cable is being moved stepped on too much, it might create unwanted flickering or discolouration. We choose the thicker ones so that they can resist wear-and-tear better.

Vanissa: How are the transparent displays are made?

Cat: GayBird wanted to use transparent displays. There are transparent TFT LCD displays available, but they are usually quite expensive. He therefore asked me to research about ways to build transparent screens. For prototyping I bought a regular computer monitor, like the ones that we use at home, and dissembled it. I realised there are different layers: a backlight unit, a TFT Array, the liquid crystal film, and some polarisers. I took away the backlight unit and the screen become transparent; however, the image became blurry and dark without the backlight.

I also got one transparent display, just for prototyping and to see its construction. They usually require a light source from the back so that the image on the screen can be seen clearly. Because GayBird wants the audience to be able to see through the screen, the light source cannot be placed directly at the back of the screen. In the current design, a LED panel is put at a 90-degree angle to the screen to act as the light source so that image on the screen can be seen clearly enough, and so that the audience can see through the screen at the same time. Each monitor is connected to an HDMI encoder. This is why each transparent screen require two power sources: one for the LED panel, and another for the encoder, which provides power to the monitor.

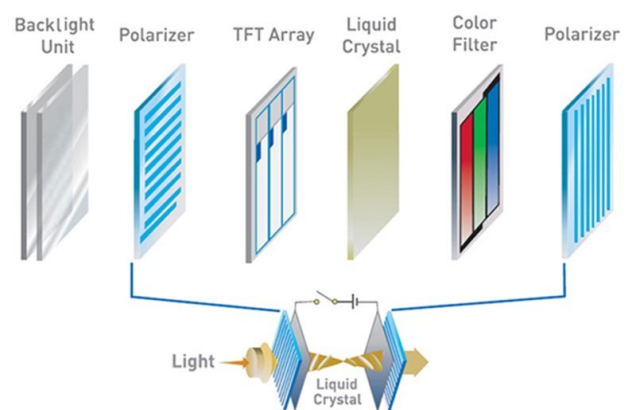


Figure 7. Typical structure of a LCD screen.

Gallery



Figure 8. Screen on the “alter” showing the back of the head of the artist.

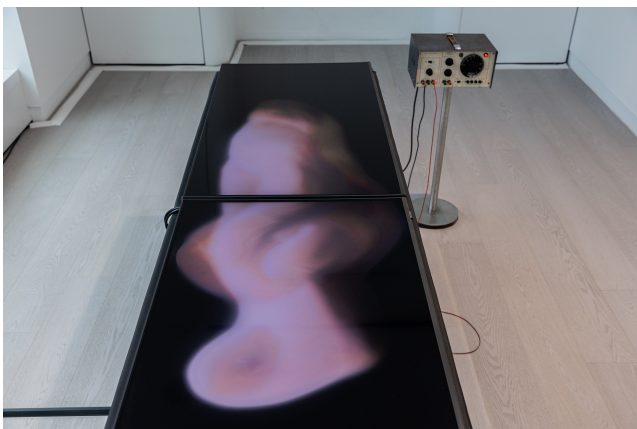


Figure 9. TV is the small room with the pre-amp of the speaker cone.



Figure 10. Transparent display in the main exhibition area showing a limb of a percussionist.



Figure 11. Side view of the main exhibition area.

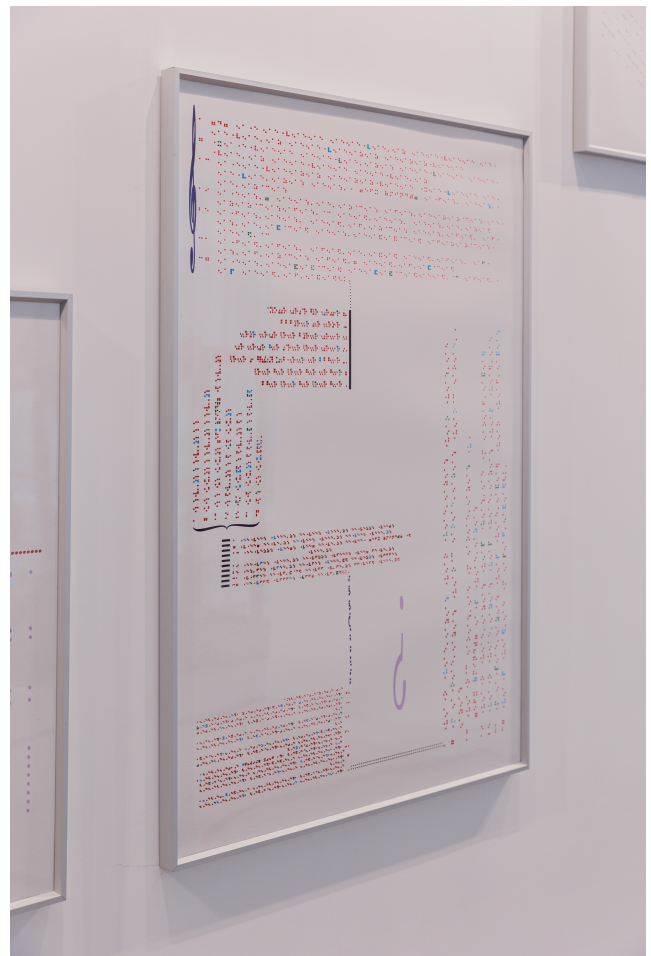


Figure 12. Artwork visualising the sound of Beethoven Symphony no.9.

Appendix

About The Missing Link

Curated by composers Gaybird Leung and Vanissa Law, The Missing Link consists of three Arts Tech exhibitions, an immersive performance, and a series of open lectures on the current issues in the use of technology in arts. By adding two essential elements—*timeliness* and *performance*— into art forms that do not originally exist in time, the two composers wish to inspire new creations and to give more perspective to existing art forms. Through the creation of multimedia artworks the artists explore ways to use technology to connect different types of media to facilitate storytelling.

The Missing Link is a two-year project funded by the Arts Capacity Development Funding Scheme (ACDFS) from 2021-2023. On top of commissioning new art works, the technical goal of this project is to develop a media synchronisation system, which will then become open source to benefit the future development of the Arts Tech industry.

About Music for 9

Music of 9 was held at H Queen's Hong Kong, from 27 April to 3 June 2023. It is a technological recreation of a youth centre in the 1990s that explores the paradigm shift of interpersonal communication over the past three decades.

Music for 9 is a media installation that fragments the video of a percussion performance, visually and aurally. Close-ups of body parts will be shown on synchronised displays, while the speakers that are paired with the displays will fill the space with impure sounds. The whole visual and audio spectrums are filled with rhythms and melodies built by re-integrated moving images and sound. Sound gradually changes from literally percussion music to sounds of the city. While the images on the screens remain unchanged, the sound that gradually breaks away from the image and took the audience to imaginations of the city soundscape.

Filled with recreational activities and new faces, youth centres were once a place to hang out, make friends and shape one's social life. Two classic tabletop games, Novuss and Table Football, were re-created in such a way that can only be controlled through a computer interface.