LUMAR: A Hybrid System for 2D and 3D Handheld Augmented Reality on Nokia S60 Series

Alex Olwal
alx@csc.kth.se

Anders Henrysson
andhe@itn.liu.se

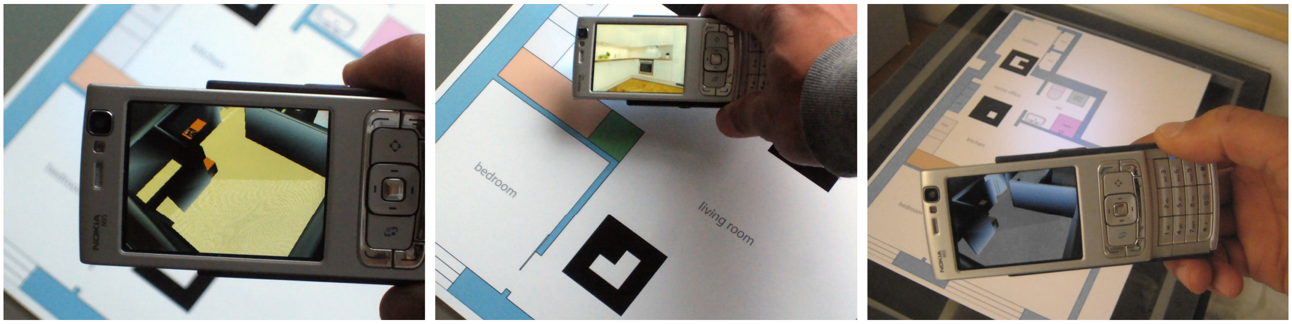


Figure 1. The LUMAR system is a hybrid system that tracks a handheld display in 2D and 3D. This allows a commercial off-the-shelf cell phone to display context-sensitive 2D information and 3D augmentations.

"Placing the phone over a room on the map will display a photograph of its interior.

When the phone is lifted, a 3D model of the apartment is superimposed and registered in 3D with the printed plan. The user can now move the phone freely to view the 3D model from an arbitrary perspective."

Interactive Apartment Exploration

One of our demo applications for the LUMAR system is an apartment ad scenario.

Floor plans are typically used to communicate the layout of rooms and their relative size in apartment ads. They are however of little use when a potential buyer wants to visualize what it is like to be inside the apartment or navigate the information associated with the different rooms, for instance.

Placing the phone over a room on the surface will load and display a photograph of the room. When the phone is moved across the surface, corresponding images are loaded automatically.

When the phone is lifted, a 3D model of the apartment is superimposed on the printed plan and registered with it in 3D. The user can now move the phone freely to view the 3D model from an arbitrary perspective. The registration makes it easy to identify which object in the 3D model corresponds to a certain graphic in the plan.

Putting the phone back on the surface switches back to 2D mode, which results in a smooth and fluid transition between 2D and 3D.

Interaction Modes

LUMAR's hybrid tracking combines the *LightSense* and *UMAR* technologies, developed by the authors, which allow fluid transition between 2D and 3D modes of interaction.

In the 2D mode, we use the phone as a tangible cursor in the physical information space that the map represents. It is convenient and intuitive to put down the phone on top of objects of interest for immediate feedback with context-sensitive information in the display.

Lifting the device, we move into 3D mode where the tracking transitions into motion-based 3D input. The information in the space can now be properly visualized with 3D graphics and overlaid on the live video feed from the camera.

The interaction modes provide fluid single-handed operation both as the device is moved within each mode as well as when it transitions between 2D and 3D modes. Buttons and joypads on the device complement the interactive aspects of each mode.

More information can be found at: www.olwal.com/?lumar

LIGHTSENSE: 2D-MODE

LightSense uses outside-in tracking to locate LED lights on cell phones. The ultra-bright LED, which normally serves as a photo light to assist taking pictures in dim lighting conditions, is here used as an easily identifiable active marker that can be detected by an external camera or photosensor. LightSense demonstrates how a combination of a computer/camera or microcontroller/photosensors can be used to locate the device on the surface. The device's distance from the surface can be estimated through the size of the detected light spot. The recovered location is continuously sent to the phone over Bluetooth.

LightSense is responsible for 2DOF position tracking on the surface. It consists of two servers running on PCs and a LightSense client on the cell phone, as shown in Figure 2 and 3. The LightSense Image Processing component detects the presence of a bright LED in the camera image. A chain of image processing filters finds the brightest spots in the image, filters out noise, and finally uses ellipse-fitting to establish the shape of the light from the LED. The center of the ellipse corresponds to the location of the phone in the camera image. The Image Processing component is implemented in C++ with a custom modular filter framework on top of the Open Source Computer Vision library, OpenCV.

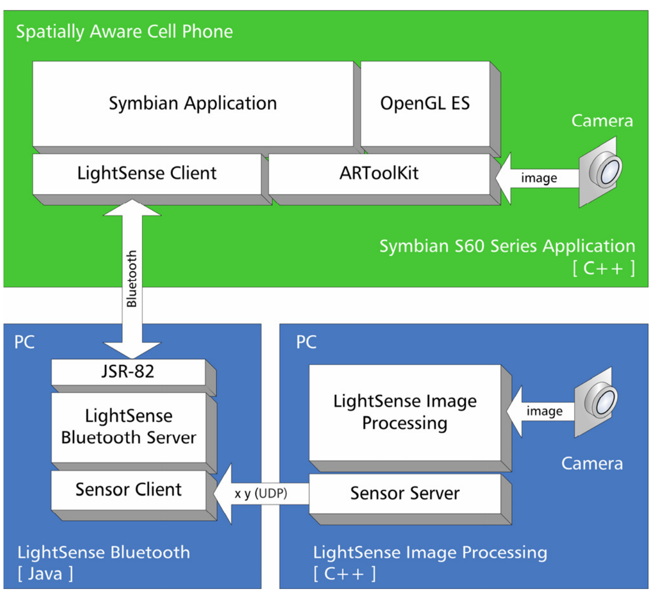


Figure 2. The LUMAR system combines egocentric and exocentric camera tracking from the UMAR and LightSense systems. While most of the work in LightSense is done with an external camera and processing on a separate PC, the UMAR tracking takes place on the device using a port of ARToolKit and OpenGL ES. This separation allows the two tracking approaches to seamlessly coexist in the system.

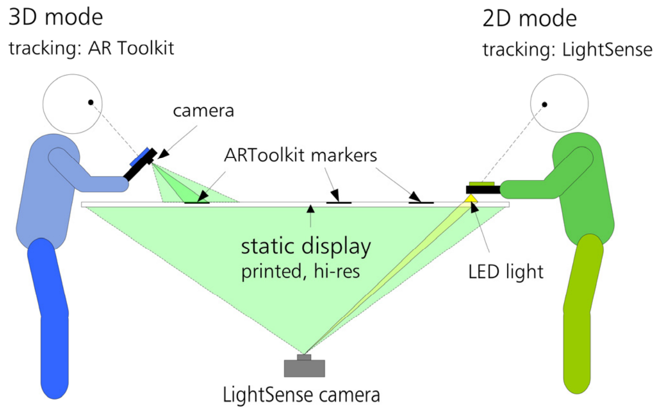


Figure 3. The LUMAR system enables fluid transitions between 2D and 3D modes. It combines the individual strengths of the respective tracking and visualization modes.

The coordinates are sent over UDP to the LightSense Bluetooth Server, a Java application that uses JSR-82, the Java Bluetooth standard, to communicate with the phone.

If the cell phone does not have a built-in controllable LED, an external LED can be attached.

UMAR: 3D-MODE

The UMAR system provides full 6DOF position and orientation tracking above the surface, where the camera is used for egocentric tracking. For this purpose, we ported ARToolKit to the Symbian platform. The performance was improved and reached interactive frame rates when the camera pose estimation algorithm was rewritten using fixed point representation. This is an essential step since most phones still lack floating point units. Floating point arithmetic is up to two order of magnitude slower when emulated in software. Further enhancements include frame-to-frame coherency thresholding to avoid expensive tracking calculations and DESP smoothing between pose estimations to stabilize tracking.

From identified markers, the algorithm uses corner and edge information to calculate the camera pose matrix relative to a coordinate system centered in the marker. The resulting matrix is used in the OpenGL ES graphics pipeline to align the physical and virtual cameras. (See Figure 2 and 3.)

More information can be found at: www.olwal.com/?lumar

Figure 4. 3D mode. A 3D model of the apartment is overlaid over the live video feed, allowing the user to view the apartment from different angles and get a good feeling for its structural properties.

APPLICATION SCENARIOS

There are many application scenarios that commonly make use of printed material, where 2D and 3D digital information could add significant and relevant information, as well as enabling the possibilities for interesting services. We described our implementation of one such scenario (See Figure 1 and 4). Additional examples where the 2D dynamic display and 3D visualizations enhance the print are outlined here.

Super map

Maps are often designed with fast navigation in mind. Thus the photometric resolution is very low with one color for each type of terrain. To make roads salient blocks of buildings are typically reduced to single objects.

2D: up-to-date aerial image data where moving the phone across the map surface controls panning. Most services associated with digital maps, such as route planning, could be supported.

3D: 3D terrain model registered with the physical map. Detailed 3D buildings with textured facades would replace the flat projections.

Information visualization

Information visualization is a field where we often like to combine high-density displays with dynamic data. Here, the print serves as an ideal grounding information space, and where the lower resolution 2D and 3D visualization can provide dynamic focus and detail.

2D: Stock prices plotted over time in a chart to give a coarse overview of their development. By moving the phone along the chart, additional data such as market index or comparable stocks can be displayed.

3D: Curve is extruded along the z-axis with the height value corresponding to trade volume.

Arena seat booking

Arena maps and folders are often used to help users select, book and find their seats. Such tangible matter is invaluable and cannot be readily replaced by solely digital information.

2D: To find a seat at an arena, move the phone over a printed section map to identify individual seats, their availability, price and number. Book a seat through a web service by simple point and click.

3D: View seats of interest highlighted in a 3D-model of the stadium. The user could fly into the model and get a feel for the view from a particular seat to ensure visibility.

Public transportation

Bus route maps often mark only a selection of the stops along the route.

2D: All individual stops can be shown along a route, and estimated time of arrival given the current traffic situation. Display geographical neighborhood map at each stop, or live video feeds from traffic cameras.

3D: A wide area 3D overview could provide current bus locations from GPS and ideal locations from time table data, to visualize delays. Color or size could be used to indicate, e.g., the current speed of a bus.

Product catalog

Printed ads are all around us trying to get our attention. They however often lack detailed product information, and to consume the product the buyer must remember the brand and model name. The Internet contains a rich set of data for various products, such as price comparisons, reviews, images, 3D models and videos. These information sources are however inaccessible from the printed media.

2D: Moving the phone over a product list, e.g., a car catalogue or restaurant menu, detailed information is shown and can be clicked for on-line price comparison services. Real-time prices for products with changing prices. Additional images and videos.

3D: Show 3D-model of the product.

More information can be found at: www.olwal.com/?lumar

IMPACT OF LAYERED INFORMATION SPACES

In visualization there is a trade-off between providing high amount of detailed information and facilitating fast navigation. In dynamic systems the user is often provided with a coarse overview of the data and means to zoom in on areas of interest. Data is often segmented into layers that can be hidden when other types of information is requested. In, e.g., Google Earth, the user is presented with a view from space from which it is possible to zoom in on any part of the globe. On the surface level, 3D terrain and buildings are added and information layers can be enabled to provide road information or photographs.

In static systems, such as printed information boards, this interactive level of detail is not possible. Static levels can be provided on separate surfaces, but this requires a switch of context with associated seams and cognitive load. In many public scenarios there is no room for more than one level of detail. In, e.g., public street maps, a high level of detail is often provided together with a legend saying: "You are here". This is a common strategy since most pedestrians want to know their current position. In contrast, public transportation maps provide a coarse overview of routes since focus is on finding the right train as quickly as possible, given a known destination. Despite rapid deployment of displays ranging in size from handheld devices to wall-mounted TV-screens, printed information will remain a dominant media given its low cost, and ease of handling and distribution.

3D information is essential for real-world navigation and conceptualization of objects. The use of the third spatial dimension further exploits human cognitive abilities to interpret visual information. Printed media used for maps and advertisement lack this capability, while stationary 2D displays require dedicated 3D interaction devices or various 2DOF modes to set the view parameters. A dedicated 3D system is however not a good option since much multimedia is inherently 2D. Browsing the web or watching a video in a 3D environment makes little sense and often degrades quality or introduces fatigue when one dimension must remain fixed.

In this work we combine the respective strengths of printed media, 2D multimedia and 3D interactive graphics. The key observation is that a mobile phone provides a high-resolution display for 2D and 3D content and enables intuitive motion-based 6DOF interaction when tracked. Modern mobile phones are able to display most media types, including videos and web pages. They provide mobile Internet access and an assortment of short range communication techniques. Instead of placing expensive and sensitive displays in public areas, we exploit the fact that most people carry a small interactive display with them at all times. To compensate for the small size of the mobile phone screen, we use the Magic lens metaphor, enabled by our exocentric tracking infrastructure. Our 6DOF egocentric tracking further enhances this setup by allowing interactive motion-based 3D navigation and interaction. It is important for the interaction techniques to blend as seamlessly as possible when transitioning between 2D and 3D mode. With our hybrid tracking strategy, we achieve this in form of motion based interaction.

The general application scenario consists of a high resolution print that gives an easily navigated overview and ensures that the tangibility of the media and its possible use as a traditional display is preserved. The high-resolution print can be important in several applications, where the superior level of detail is vital, yet impossible to reproduce with a digital display. Subsequent 2D information layers are digital and viewed by placing the phone on the surface of the print. They can be either continuous, providing higher information density than the print for the same region of interest, or they can be segmented into areas, each associated with various kinds of 2D multimedia, such as videos or web pages. When the phone is lifted, a transition occurs into 3D mode, where 3D information serves as an augmentation of the print or as a continuous 2D layer now registered in 3D. Thus there is a close spatial relationship between 2D and 3D information. The phone acts as a looking-glass that magnifies the information content.

CONCLUSIONS AND FUTURE WORK

We have demonstrated a hybrid tracking system that combines printed media with digital 2D multimedia and 3D graphics, using handheld AR. This is important since all three media types have their individual strengths and together contribute to empower our human perception. The single-handed interaction is based on motion in both 2D and 3D modes which minimizes the interaction seams.

2D and 3D data could be downloaded to the phone on the fly, which would require service discovery and data transfer. Since we know beforehand what the print looks like we could also provide feature data in addition to 2D and 3D information. That could allow us to replace or extend the ARToolKit tracking with natural feature tracking, resulting in potentially even smoother transitions between 2D and 3D modes.

After our implementation and exploration of three-layered information spaces, we find it interesting to in future work explore other possible application scenarios, as well as their associated interaction techniques.

More information can be found at: www.olwal.com/?lumar

INSTALLATION ON NOKIA N95 8GB

1. Send the LUMAR.sis over Bluetooth to the phone

2. Open the received message and install the application.

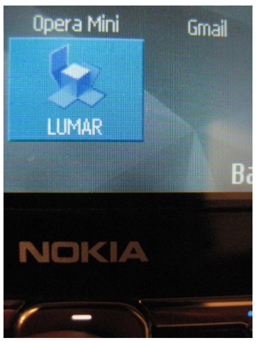
TESTING LUMAR'S 3D AUGMENTED REALITY

Since LUMAR relies on some additional PC-side software and sensing for its 2D-part, only the 3D mode can be tested in the submitted standalone application, since there were no facilities in the submission system to submit additional material (except the movie and this PDF).

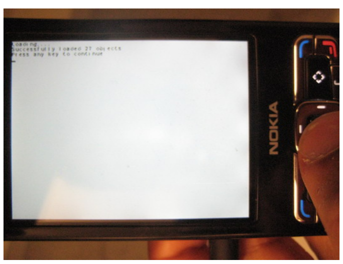
To test the 3D Augmented Reality:

1. Print out the map with the markers (attached as 4 separate pages at the end of this document) – at least one visible markers is needed in the camera view at all times.

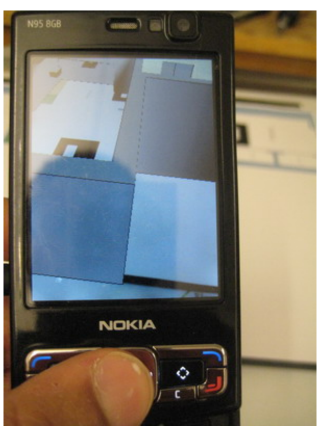
2. Start the LUMAR application



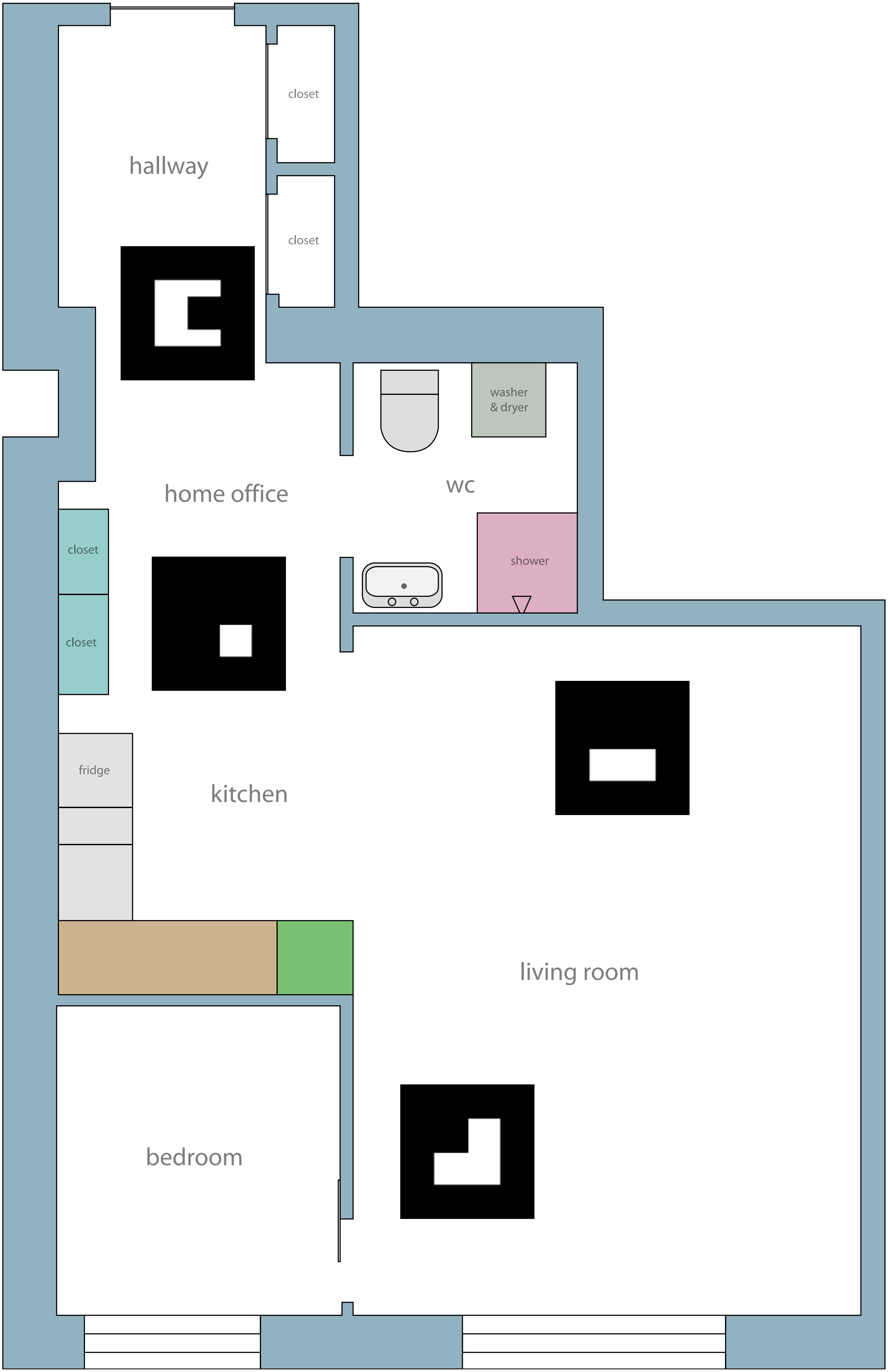
3. The application is starting up…



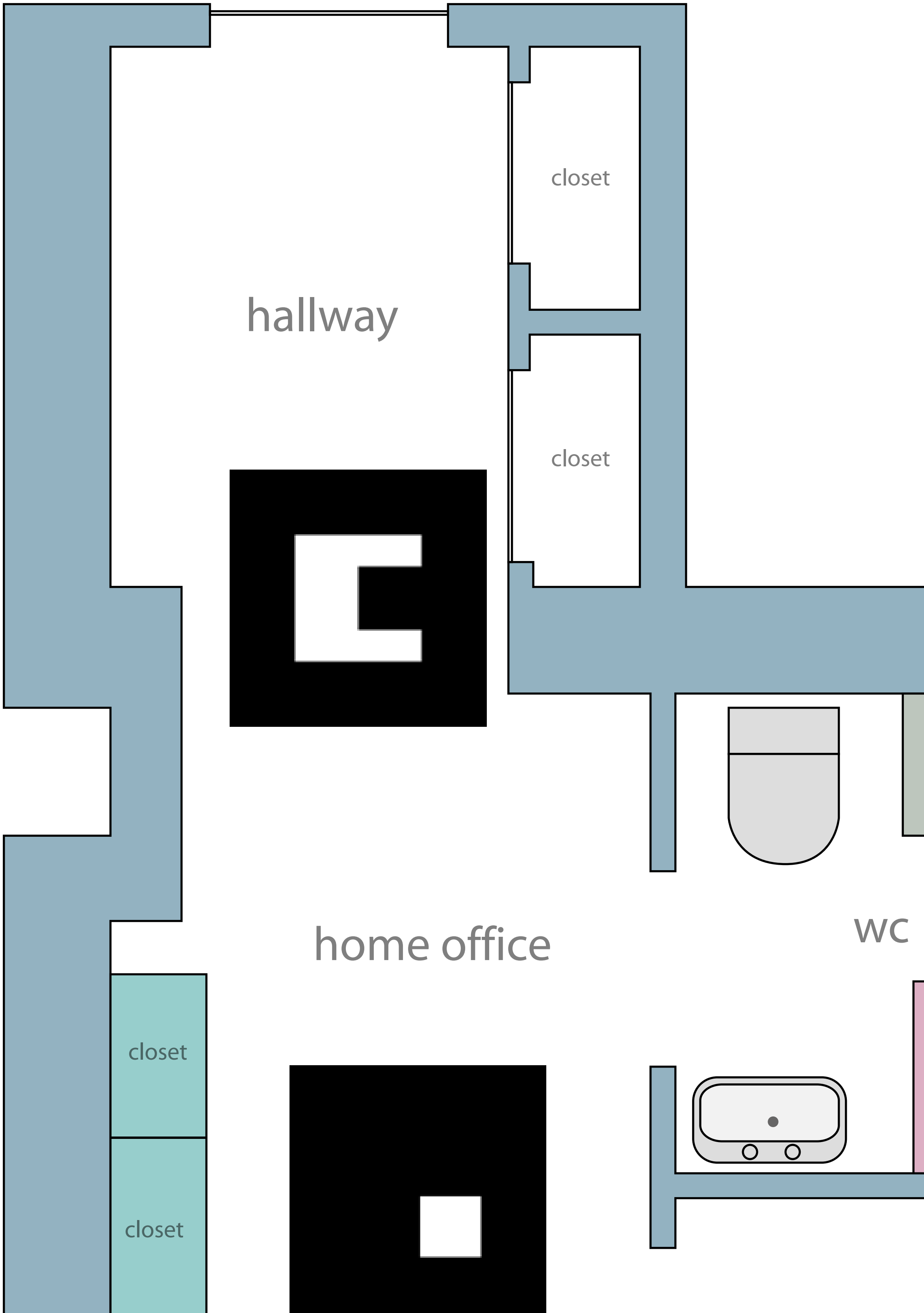
4. …and when loaded, point the phone's camera towards one of the printed markers. A 3D model should now appear registered with the video, and the perspective of the overlaid model will change naturally as the phone is moved relative the real world.

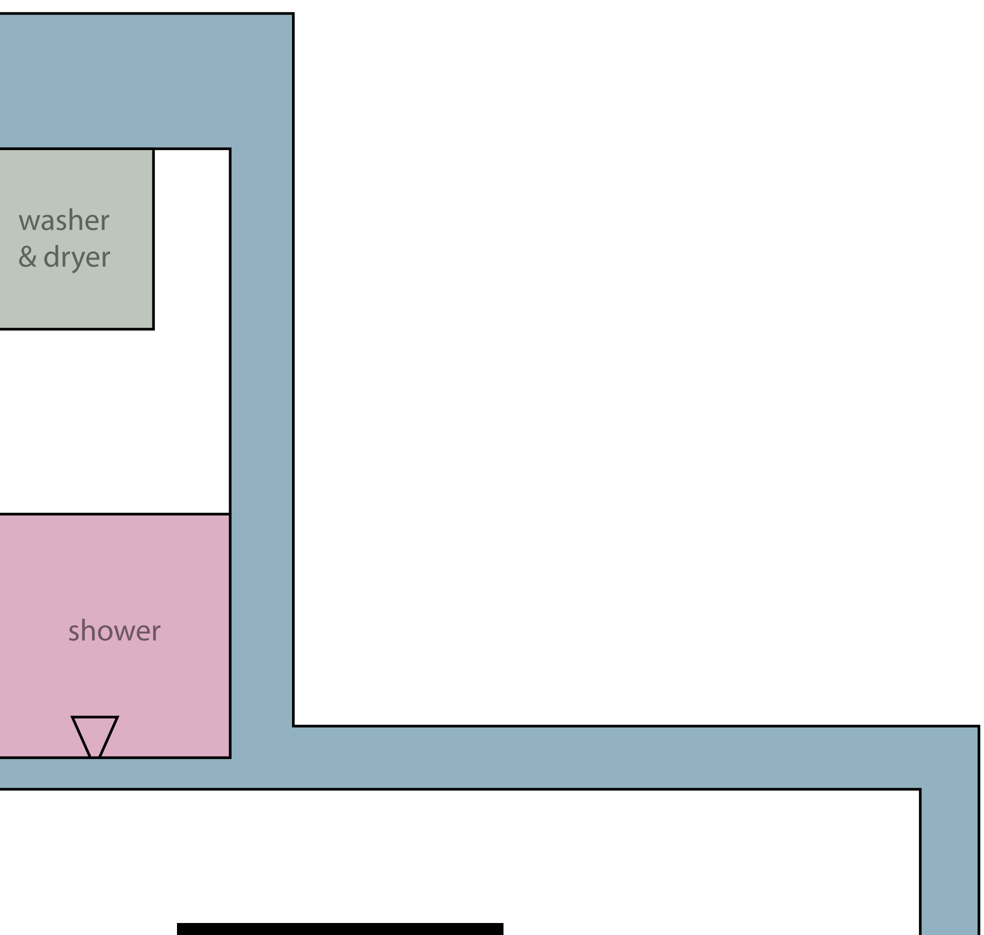


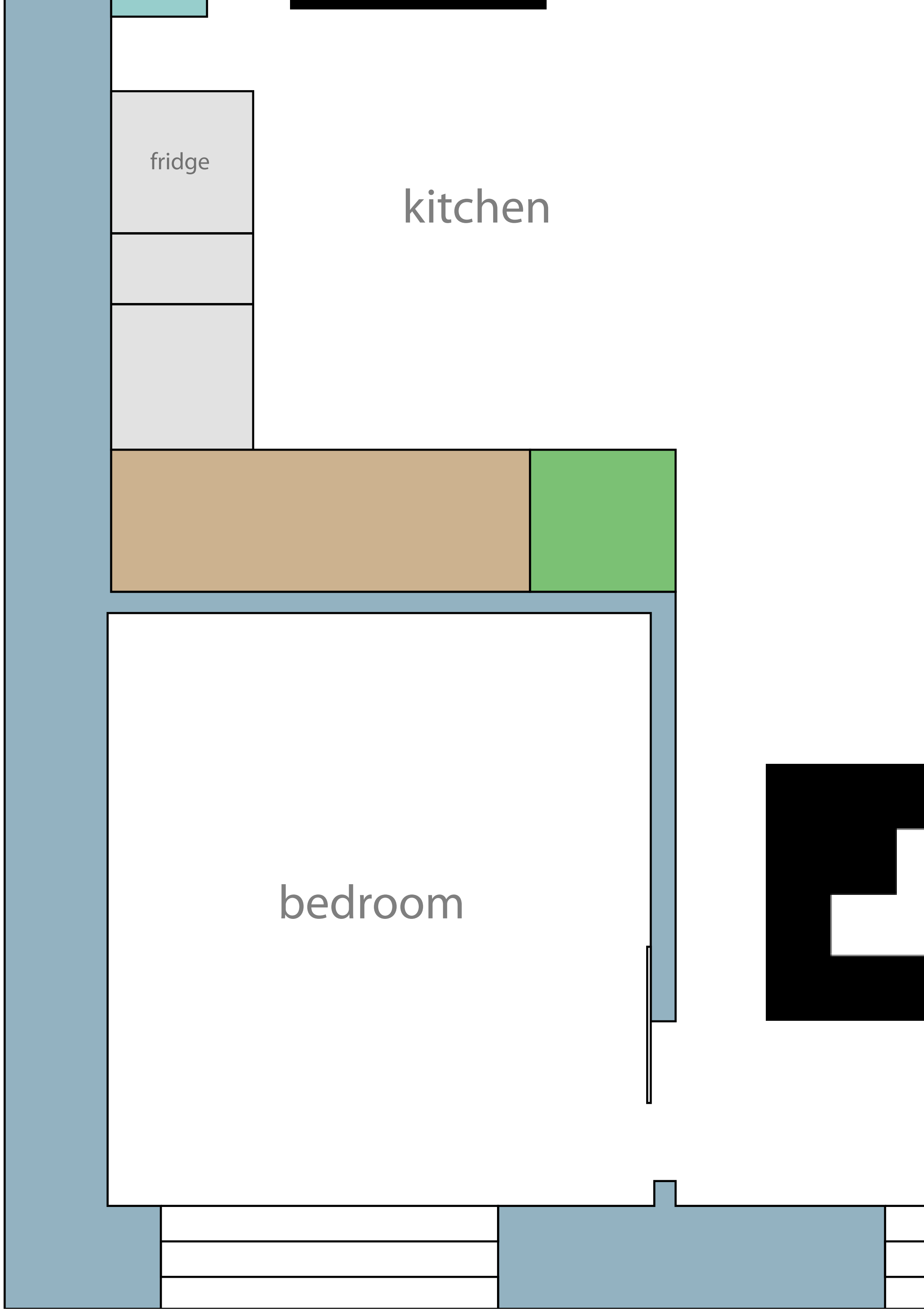
More information can be found at: www.olwal.com/?lumar

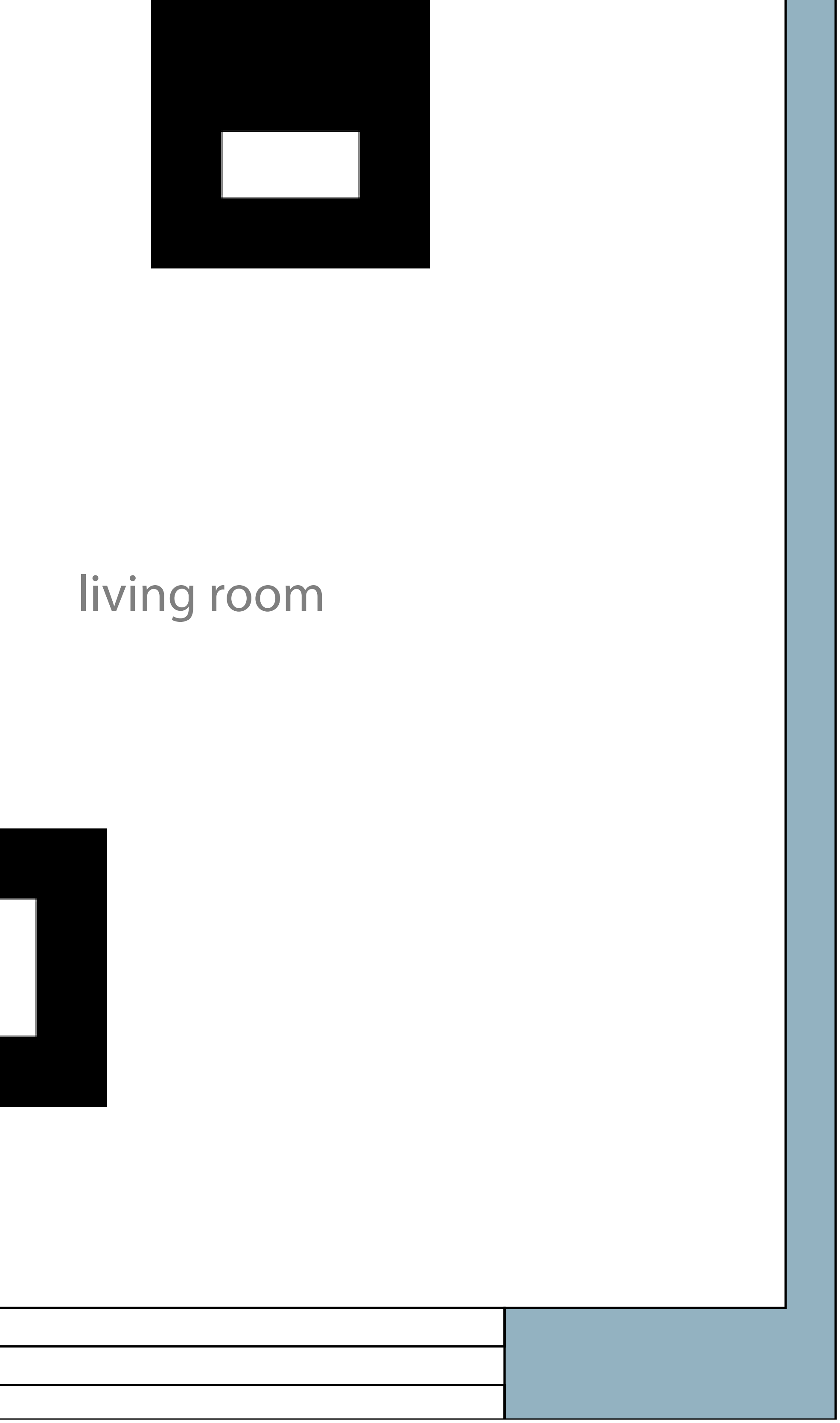


closet
hallway
closet
washer
& dryer
home office
wc
closet
shower
closet
fridge
kitchen
living room
bedroom


closet
hallway
closet
home office
wc
closet
closet


washer
& dryer
shower


fridge
kitchen
bedroom


living room