Implementation of a new text-video encoding-decoding mechanism using morse code

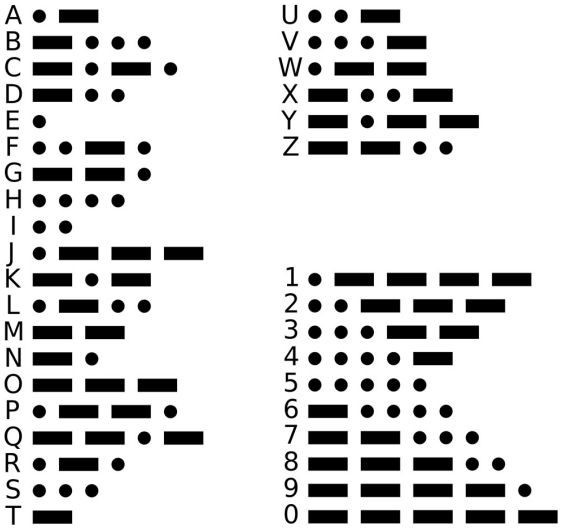
Step 1:

Finding a phenomenon and a question to ask about it:
Morse Code is something that intrigues me because of its simplicity and usefulness. It is used for transmitting messages in which letters of the alphabet and numbers are represented by various sequences of written dots and dashes, or short and long signals such as electric tones or voltages. Such patterns of short and long signals can be simulated by something as simple as flashing of lights. The questions addressed here are (1) Taking text data as input, how can we encode it into video data using morse code patterns? (2) Given such video data, can we identify morse code patterns and decode them efficiently?

Step 2:

Understanding the state of the art:

This is how some characters are represented in morse code using dots and dashes:



Several computer vision models have already been implemented to decode morse from known patterns. This thesis implements morse decoding from finger gestures. Another implementation of morse decoding can be found here, working on morse input generated by blinking of eyes. Instead of aiming for advanced approaches, we could

start off with a baseline approach of using flashing LEDs for the video part. For the rest of this discussion, we will assume that the video being referred to here is that of an LED flashing in a particular pattern and the time scale to be in milli-seconds, for the sake of simplicity.

Step 3:

Determining the basic ingredients:
For the encoding task, the input can be a simple text phrase and the output would be a video of an LED flashing in the corresponding morse code pattern. The length of the text could be limited to a few words, say 20. There will have to be a fixed representation of dots and dashes in LED flashing patterns. As a general rule of thumb, dashes should be 3 times longer than dots and timing between dots and dashes should be the length of a dot. For the sake of simplicity, we will fix the representation of a dot as the LED being lit for one milli-second, dash as the LED being lit for 3 milli-seconds, the space between letters of a word as LED being off for 3 milli-seconds and space between words as LED being off for 7 milli-seconds.

Step 4:

Formulating specific mathematically defined hypothesis:
We hypothesize that given a representation of morse code parameters, it is possible to encode text data to video and the other way round as well.

Step 5:

Selecting the Toolkit:
The encoding from text to video can be performed in a fairly simple manner using simple functions. Decoding video input to text would be a challenging task if the pattern is not already known. Predefined template matching algorithms for pattern recognition could be employed here.
Spiking Neural Networks coupled with input from Dynamic Vision Sensors could also be considered as a method for implementing the video to text decoding functionality.

Step 6:

Planning the model:
The text could be converted to a character-wise representation as morse code. After this, an image frame (or a set of frames) per morse character could be generated which could then be appended to form a video. In our simple base case, the image frames could be an on or off LED.

For video data, we could first convert it into image frames and then recognise morse patterns through template matching algorithms. These patterns could be converted to morse representations which could further be converted to text inputs. For simplicity, we would only output the top five text representations of morse patterns.
Further discussion about using spiking neural networks and dynamic vision sensors could prove to be useful in video to text decoding.