Accelerating Multimedia Applications using the Intel SSE 4 and AVX ISA

Motivation: A great portion of the multimedia applications such as template matching, object detection and face recognition involves intense work on searching, comparing and mathematical computation. Given the facts that high resolution images and HD videos are widely adapted, and enormous information is stored in the database as the source of recognition, multimedia applications need to be continuously accelerated to catch the pace of the changes on the resources side. Besides the efforts spent by the algorithm designers, recent innovations in computer architecture together with their newly defined Instruction Set Architectures (ISA) give us the opportunity to accelerate the multimedia applications from the standpoint of hardware/software implementation.

The Intel SSE 4 and AVX ISAs: The Intel SSE 4 and AVX will use the 128-bit XMM or 256-bit YMM registers respectively to perform searching, comparing and vector computation. For example, the 128-bit register can store 16 “char” type data since the size of a “char” is 8 bits. Then by using the SSE4.2 string and text ISA when searching for a certain character among 16 characters, one instruction is needed instead of comparing with 1 “char” at a time and using 16 instructions to complete this task as what is normally implemented. The 256-bit YMM registers are an extension to the 128-bit XMM registers, which is used by the AVX ISA to accelerate the vector computation. The first commercial CPU product that supports SSE4.2 was released around 2009 while the first commercial CPU product that supports AVX was released in 2011. Currently the fact is that existing compilers are far from being eligible to fully optimize the program using these new ISAs. In this case, the major goal of this project is to re-write the core part of these multimedia applications by using the SSE4 and AVX ISAs for runtime speedup. Besides this, I would like to explore the intrinsic of automatic compilation based on the two ISAs.

The Intel Opencv Library: The OpenCV (Open Source Computer Vision) is a library of programming functions for real time computer vision. It contains many optimized functions that cover different multimedia applications ranging from image processing to machine learning. The source codes are well written in multiple computer languages. For the concern of performance, this project will work on modifying the Cpp source code. The applications focused in this project include object detection and face recognition.

Performance measurement: One of the major reasons of choosing applications like object detection and face recognition, is that after acceleration, hopefully the speedup can be directly shown visually. If the acceleration is not that prominent, profiling will be done to offer more accurate statistics for the speedup. Besides the measurement of application runtime, the memory usage is another major measure. The best scenario is that a small portion of memory is consumed to exchange for a significant speedup.

Tentative Project Plan: 2 weeks: getting familiar with the applications, ISAs (write small examples). 4 weeks: modify the codes, make test from real inputs. 1 week: summary and report