An Integrated Memory Array Processor Architecture for Embedded Image Recognition Systems

Shorin Kyo, NEC Corporation Shin'ichiro Okazaki, NEC Corporation Tamio Arai, University of Tokyo Conferência ISCA 2005,

Resumo por: Leandro Rodrigues Magalhães de Marco RA 009089

Existem muitos esforços para desenvolver processadores embarcados para aplicações multimídia, com performance suficiente e baixos custos (que envolvem também baixo consumo de energia). O artigo descreve o IMAP, que é um processador SIMD (Single Instruction, Multiple Data), que se propõe a ser um processador embarcado para sistemas de reconhecimento de imagem.

Tarefas de reconhecimento de imagens de vídeo, assim como aplicações multimídia, possuem um alto nível de pararelismo e necessitam de resposta em tempo real. Entretanto algoritmos de reconhecimento de imagens tem uma diversidade muito maior, o que torna necessário além de performance, flexibilidade.

Devido a essa necessidade por flexibilidade, foi adotada a estratégia de desenvolver técnicas de paralelização que permitam mapear diversas tarefas de reconhecimento de imagens em um processador SIMD altamente paralelo.

Para facilitar a técnica de paralelização, foram definidos sete grupos de operações que por sua vez foram classificadas em quatro categorias de acesso à memória. Esta classificação levou em conta critérios de localidade e confinamento da seqüencia de atualização de pixels. Com esses padrões em mente, foi feita a escolha de uma arquitetura em que cada PE (elemento de processamento) possuí uma quantidade de memória endereçável, e está conectado aos outros PEs em anel. Nesta arquitetura quatro técnicas de paralelização foram definidas. Cada uma para atender uma das categorias de acesso a memória.

Para uma implementação eficiente, e para utilizar um alto nível de paralelismo no IMAP, uma extensão da linguagem C, chamada C unidimensional (One Dimensional C, IDC) foi utilizada. No IDC, entidades associadas ao vetor de PEs são declaradas com a palavra-chave sep (ou separate). Quando o sep é utilizado, são denotadas operações explicitamente paralelas. Ou seja, diversos PEs utilizarão aquela entidade ao mesmo tempo.

O IMAP é constituído de um procesador de controle (CP), que é um RISC de 16 bits de propósito geral com um pipeline de 6 estágios, um vetor de 128 PEs (16 grupos de 8 PEs). Cada PE é um RISC de 8 bits com um pipeline de 3 estágios. O CP possuí 32 Kb de cache de programas e 2 Kb de cache de dados. Cada PE acessa 2 Kb de memória (IMEM). A soma de todas as IMEMs (256 Kb) é o espaço de trabalho dos PEs. Entidades declaradas com sep, ficam armazenadas na IMEM e são mapeadas na EMEM de 64 M que além de servir de swap para a IMEM, também é cache de programa e dados para o CP. O consumo de energia do IMAP é de em média 2 Watts.

Os resultados da utilização do IMAP de 100 MHz, comparado a um Pentium 4 de 2.4 GHz, utilizando códigos IPC para o IMAP e códigos seqüenciais para o Pentium 4 mostram um ganho proporcional ao nível de paralelismo esperado. Em todos os casos O IMAP supera o Pentium 4. Quando comparado também a códigos com otimizações MMX, códigos IDC rodando no IMAP são 3 vezes mais rápidos do que códigos MMX e 8 vezes mais rápidos do que código C sem otimizações. Foi feito um benchmark com uma aplicação de detecção de veículos. Este revelou dois casos em que o Pentium 4 se sai melhor que o IMAP. Em um dos casos o paralelismo existente era modesto comparado ao númeor de PEs. No outro caso, um processamento seqüencial era necessário.

Pelos resultados pode-se constatar a performance alcançada pelo IMAP. Um IMAP de 100 MHz, consumindo apenas 2 Watts, oferece uma performance em média quatro vezes maior que a de un Pentium 4 de 2.4 GHz que consome cerca de 100 Watts. Entretanto, o benchmark feito com uma aplicação do mundo real mostrou que em alguns casos o fato da arquitetura ser puramente SIMD inibe um uso mais eficiente. Direções futuras apontadas envolvem melhorias de performance e maior flexibilidade no controle do vetor de PEs.