



Skräddarsydda systemkretsar avstyr krocken

Storlek, effektivitet, flexibilitet och pris avgör genomslaget



Ju fler bilar som utrustas med aktiva säkerhetssystem som hjälper föraren att köra, desto färre kommer att dödas på våra vägar i framtiden. Det europeiska bedömningsystemet NCAP är en viktig marknadsfaktor för att nå detta mål, överkomliga priser på stödsystemen är en annan. För att kamerabaserade ADAS-system (Advanced Driver Assistance Systems) ska nå stora volymer måste modulerna bli billigare och den totala materialkostnaden minska. Samtidigt förblir effektförbrukningen en nyckelfaktor.

Dagens kameramoduler innehåller flera olika beräkningsenheter på kretskortet, såsom dedikerade hårdvaruacceleratorer, FPGA:er, DSP:er och processorer. Här presenterar vi en lösning på en programmerbar energisnål kameraplattform.

DE OLIKA STEGEN VID BILBEHANDLING har sina särdrag (figur 1). Först hanteras sensordata genom exempelvis filtrering eller bildförbättring. Varje pixel bearbetas och



Av Michael Staudenmaier & Stephan Herrmann, Freescale

Michael Staudenmaier började på Freescale 1999 och har idag ansvar för analys och val av acceleratorarkitektur till ett förarstödsystem för fordon (ADAS). Tidigare har han bland annat ansvarat för definition och utveckling av multimedia-IP i informationssystem för fordon. Hans huvudområde är hur datorgrafik används i fordonstillämpningar.



Stephan Herrmann leder arbetet att definiera arkitekturen hos video- och bildsystemen i Freescales nästa generation ADAS-system. Han har jobbat på Freescale sedan 2005, till en början med att utveckla kamerabaserade ADAS-system. Dessförinnan har Stephan bland annat forskat inom bildanalys och dess implementering i hård- och mjukvara. Han doktorerade vid Munich University of Technology år 2005.

antalet operationer per sekund är enormt, men dataflödet kan organiseras med hjälp av relativt små lokala minnen.

I steg två extraheras bildfragment, till exempel kanter eller hörn från bilden. Även här arbetar man på pixelnivå i ett glidande fönster och har liknande möjligheter till parallelliserbarhet som i förbehandlingssteget. De extraherade fragmenten är antingen baserade på pixlar eller på komprimerade deskriptorer i sammansatta datastrukturer.

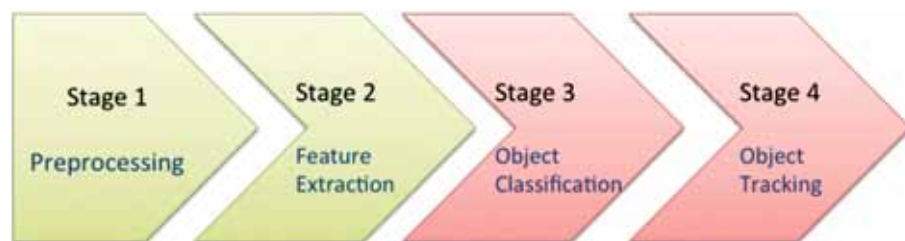
I det tredje steget klassificeras objekt utifrån de sammanställda bildfragmenten. Här minskar datamängden radikalt. Algoritmerna för att detektera objekt har invecklade styrflöden och accessar data i olika delar

av minnet. Kravet på bearbetningsprestanda är högt även här eftersom det är många bildfragment som ska sammanställas och jämföras mot en databas av klassificerade objekt, som måste vara omfattande för att bli robust och för att minimera antalet felklassificeringar.

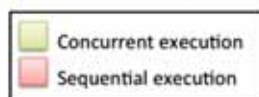
I det fjärde steget följs de intressanta objekten över ett antal bildrutor, och det konstrueras en omfattande modell av miljön runt bilen, varefter till slut situationen bedöms för att trigga eventuella åtgärder. Bearbetningen är återigen kraftigt icke-linjär men datamängden är ganska liten.

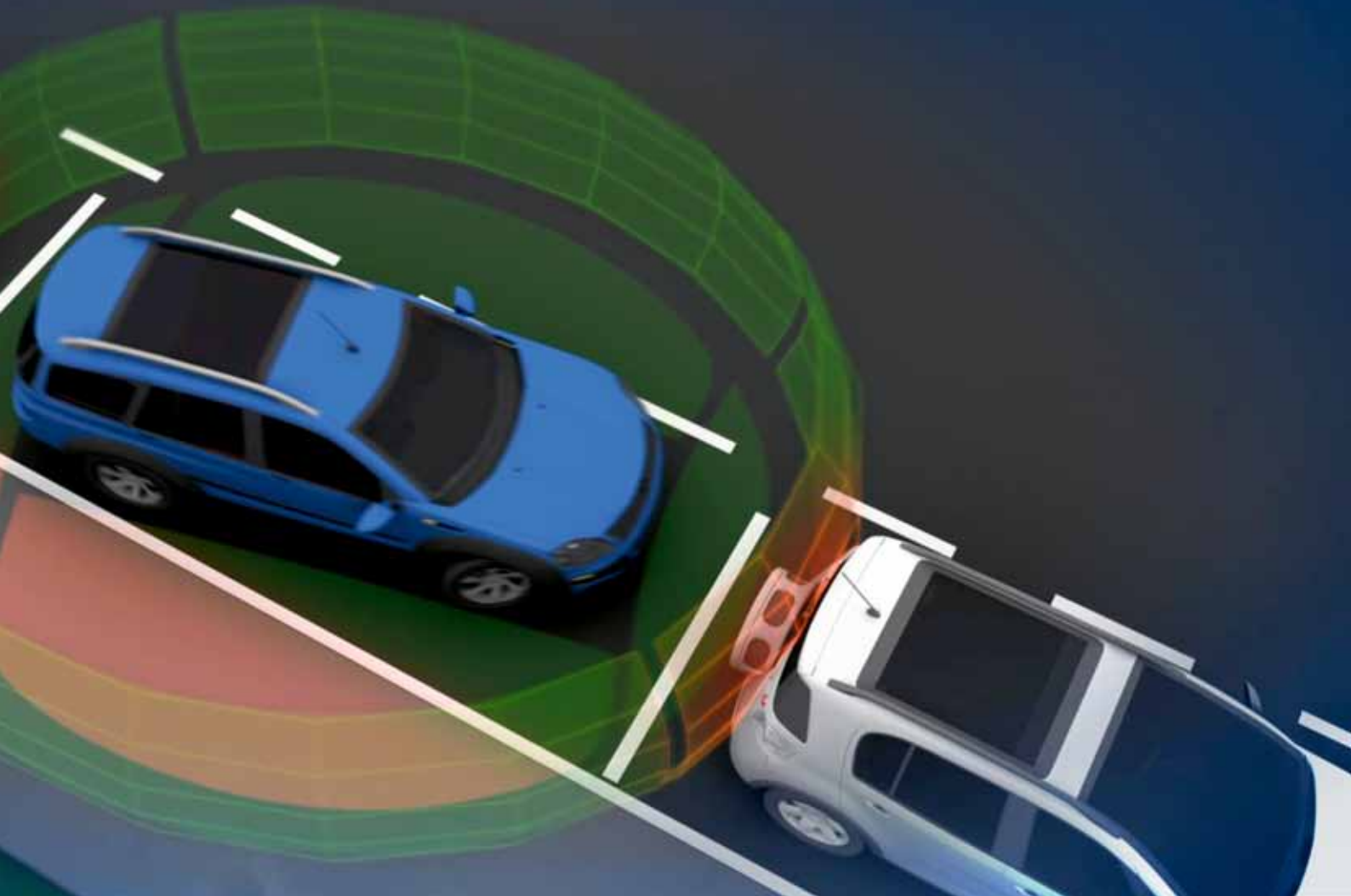
Alla nämnda bildbehandlingssteg måste vara flexibla och programmerbara för att snabbt kunna anpassas till förändringar i algoritmer och marknadskrav. De kräver också olika beräkningsarkitekturer för att ha bra prestanda och låg effektförbrukning.

DEDIKERADE ACCELERATORER ger högst prestanda till lägst effektförbrukning (figur 2, se Application Specific Signal Processor Core och Physically Optimized Macro). Arkitekturerna består vanligen av en uppsätt-



Figur 1. De fyra stegen vid bildbehandling.



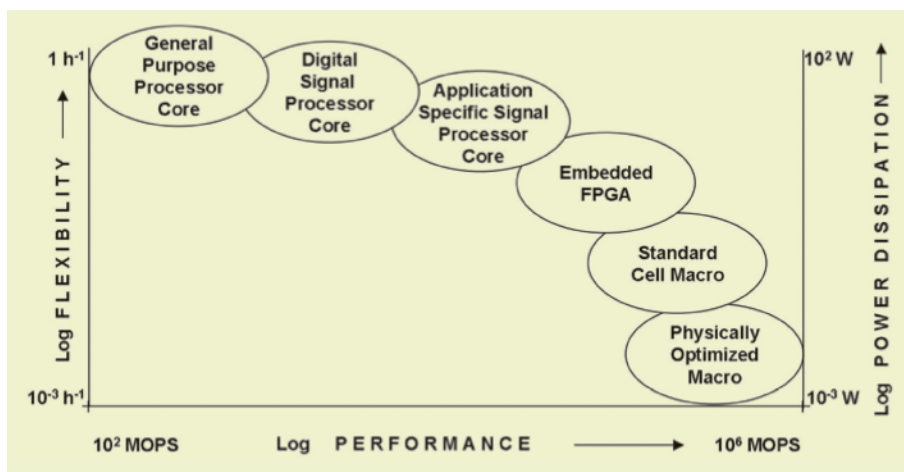


ning hårdvarumoduler som kan utföra en specifik bildbehandlingsuppgift. De är implementerade som ett delsystem och kan hantera hög bandbredd. Eftersom funktionen vanligtvis är skräddarsydd är möjligheten till återanvändning mycket begränsad.

TYPISKA FUNKTIONSBLOCK ÄR:

- FIR-filter med parametriserade filter-kärnor
 - skalningsfilter som skapar en pyramid av allt mindre versioner av den högupplösta ingångsbilden
 - Sobel-filter (figur 3)
- Histogramberäkning för bildstatistik
- Rörelseanalys som använder särskilda datatyper (t ex baserade på bildfragment) eller regelbundenheter på högre nivå (t ex blockmatching)

Dedikerade acceleratorer är alltid effektivast vad gäller pris och prestanda. Däremot är de inte flexibla och är därför lämpade för algoritmer som är fasta (inte behöver förändras). Detta gäller för det första steget i bildbehandlingen där funktionerna ofta implementeras i en rörledning genom vilken data strömmas genom behandlingsenheterna. Således adderar de enbart liten fördröjning och fungerar med lokala minnen.



Figur 2. Flexibilitet kontra prestanda hos olika arkitekturer.

FPGA:ER ÄR MYCKET FLEXIBLA och fungerar bra för algoritmer med en hög grad av inbyggd parallellism, som pixelbaserad bearbetning. Därför passar de bra till steg ett och två i bildbearbetningen. En utmaning är dock att nå funktionssäkerhet, eftersom en FPGA kan omkonfigureras under drift.

FPGA:er har i allmänhet också hög effektivförbrukning på grund av sin fysiska uppbyggnad med en väsentlig mängd outnyttjade förbindelser. Denna struktur gör

också att produktionskostnaden ökar (fler masklager) vilket är en utmaning när volymerna går upp.

DIGITALA SIGNALPROCESSORER erbjuder en bra kompromiss för algoritmer med parallell pixelbearbetning liksom vid bearbetning av kundspecifika datastrukturer eller komplexa styrflöden. Det gör DSP:er lämpade för steg två och tre i bildbearbetningen.



Figur 3. Sobel-filter (vänster) och Harris Corner Detector (höger) används vid bildbehandling för att detektera kanter och hörn.



DSP:er med VLIW-arkitektur har idag hög komplexitet och långa rörledningar. Därmed blir antalet kärnor i en DSP-arkitektur mycket mer begränsad än i arkitekturer med enklare kärnor.

DSP:er skalade bra så länge som frekvensen ökade med varje ny generation, men eftersom detta inte längre gäller så kommer en ensam DSP inte att räcka till i många framtida tillämpningar. DSP-baserade lösningar kommer att ha begränsad skalbarhet då det ännu inte existerar en etablerad programmeringsmodell för flerkärniga DSP-tillämpningar.

Samtidigt har signalprocessorerna fått stark konkurrens från riscarkitekturer, vars bearbetningskapacitet för parallella operationer närmar sig DSP:ns med varje ny generation genom integrerade hjälpprocessorer med SIMD-arkitekturer (single instruction multiple data). DSP:er erbjuder alltjämnt fördelar jämfört med riscarkitekturer genom att de använder den så kallade Harvardarkitekturen med multipla databussar som ger högre datagenomströmning. Både

riscprocessorer och DSP:er levererar vanligtvis några tusen Mips (miljoner instruktioner per sekund) vilket är åtminstone en magnitud mindre än en dedikerad accelerator eller FPGA kan uppnå.

GRAFIKPROCESSORER, GPU:ER, faller inom kategorin applikationsspecifika processorer (ASIP:er). De erbjuder vanligtvis väsentligt lägre parallell bearbetning än dedikerade accelerators eller FPGA:er och instruktionsuppsättningen är mindre flexibel än hos DSP:er. GPU:er passar för delar av steg två och tre i bildbearbetningen. Tyvärr stöder de flyttal, grafik och annat som inte behövs i denna typ av plattform utan endast slukar effekt och adderar onödig kostnad. GPU:er är därför inte lämpade för inbyggd bildbearbetning, även om de passar väl för snabb algoritmutveckling på en pc.

MASSIVT PARALLELLA SYSTEM är en annan typ av ASIP:er som passar väl för den beräkningsintensiva pixelbearbetningen i steg ett och två i bildbearbetningen. Programmerbarheten hos denna typ av system

ger stor flexibilitet samtidigt som funktionaliteten kan definieras genom programvara.

I motsats till grafikprocessorer kan en massivt parallell bildanalysaccelerator optimeras för specifika funktioner, såsom optimering av beräkningsenheter (computation units, CU) vilket är motsvarigheten till shaders i grafikprocessorer. Vid bildbehandling finns egentligen inga krav på flyttalsfunktionalitet, utan heltalsaremetik fungerar bra. Beräkningsenheterna behöver således inte ha flyttalsenheter, vilket minskar kiselytan väsentligt.

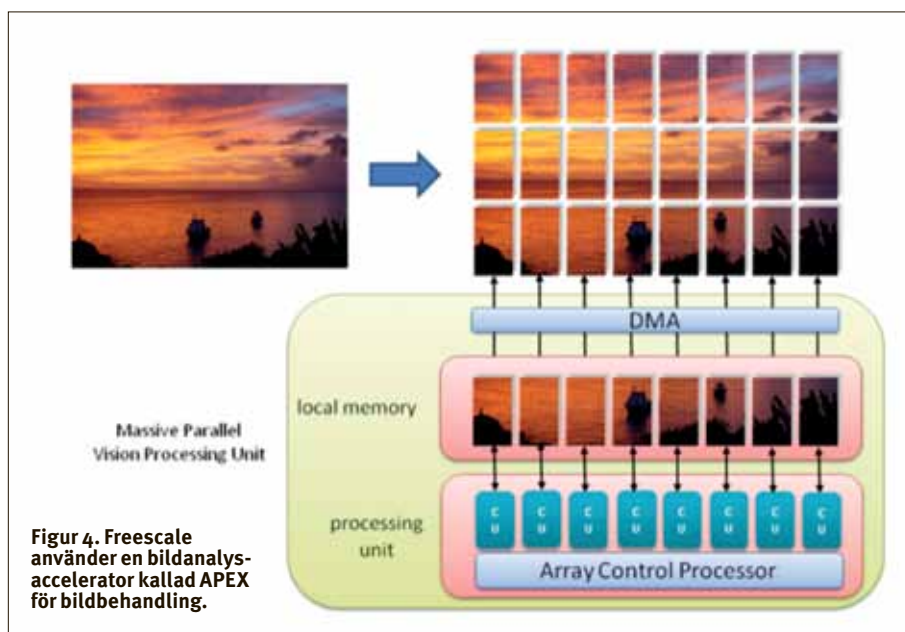
Dataflödet och minnesåtkomsten måste också optimeras. Minnesåtkomsten vid pixelbearbetning är oftast förutsägbar och linjär. Det gör att den kan optimeras för en specifik funktion och även ta hjälp av speciell hårdvara som DMA (Direct Memory Access). Denna optimering resulterar i en ASIP för bildbehandling som både blir liten och strömsnål.

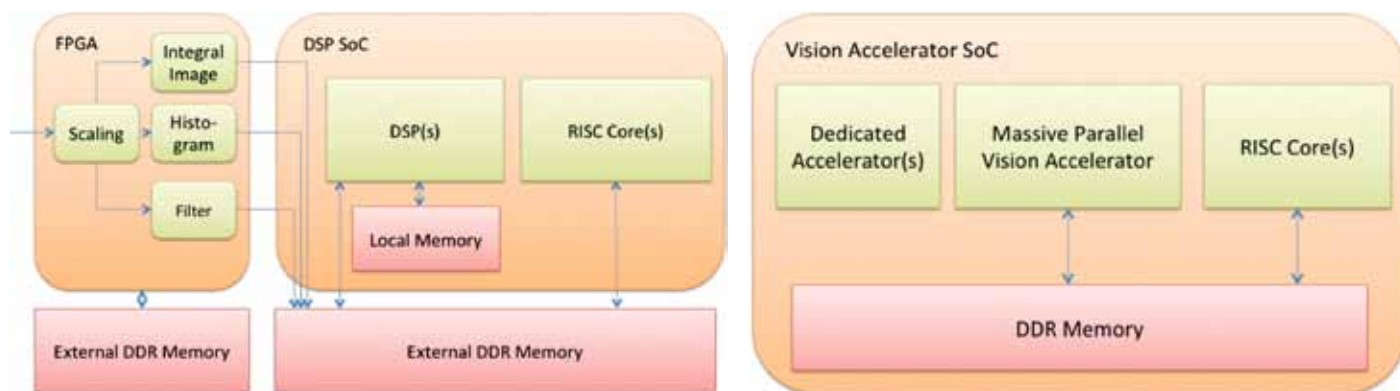
Freescale erbjuder en bildbehandlingsenhet för bakåtriktade bilkameror (rear-view camera) som använder en parallell bildanalysaccelerator kallad APEX (figur 4).

EN KOMPLETT SYSTEMLÖSNING måste kombinera flera olika arkitekturer för att bli så energi- och kostnadseffektiv som möjligt. En eller flera riscarkärnor används exempelvis alltid för steg fyra i bildbearbetningen samt för operativsystemet och applikationsprogramvaran.

I dagens ADAS-system är det vanligt med hybridlösningar som kombinerar FPGA:er, DSP:er och riscarkärnor på samma kretskort. Därtill kommer exempelvis DDR-minnen med ett stort antal anslutningar. Resultatet är att kostnaden, storleken och effektförbrukningen skjuter i höjden. Figur 5 visar exempel på en sådan hybridlösning, där steg ett av bildbehandlingen hanteras i en FPGA och steg två delas mellan FPGA:n och en DSP.

I framtida kamerasystem, som ska säljas i stora volymer och nå högt betyg i Euro NCAP-tester, är det ett starkt tryck från marknaden att använda systemkretsar som inkluderar all signalbehandling och en stor





Figur 5 och 6. En hybridlösning (vänster) med både FPGA och DSP blir dyrare och mindre energieffektiv än en optimerad integrerad arkitektur (höger).

mängd minne. Ett billigare och energieffektivare alternativ till lösningen som visas i figur 5 är att implementera de fasta delarna av steg ett i bildbehandlingen – såsom skalning, histogram och FIR-filter – i dedikerade accelerators, se figur 6. Resten av pixelbearbetningen i steg ett och två bör hanteras av en massivt parallell bildanalysaccelerator, som APEX, istället för en DSP. Fördelarna är lägre energiförbrukning samt en programmeringsmodell som stöder obegränsad parallellism. Likaså kan en riscjärna med DSP-extension hantera de övriga bildbehandlingsuppgifterna som DSP:n utför i den diskuterade hybridlösningen ovan utan att effektiviteten minskar i någon större utsträckning.

PROGRAMVARAN I ADAS-SYSTEM är en dyr post eftersom den är extremt komplex och kräver omfattande tester och kvalificering för att uppfylla alla krav på säkerhet och tillförlitlighet. Därför är ökad återanvändning viktigt.

Idag krävs det flera olika steg i arbetet att utveckla programvaran. Det första är att gå från datormiljön där algoritmen utvecklas till den inbyggda tillämpningen.

De flesta systemleverantörer använder olika inbyggnadsplattformar parallellt, vil-

”Genom att utföra flera bildbehandlingssteg utan att belasta externa minnen med delresultat kan man skapa betydligt mer energi- och kostnads-effektiva lösningar”

ket ofta kräver portering till en annan arkitektur. Eftersom programvaran är mycket hårdvaruberoende – särskilt i de två första bildbehandlingsstegen – är det avgörande hur den porteras för att den senare ska kunna återanvändas.

Lösningen är att använda standardiserade programmeringsmodeller för parallell bearbetning och att tillhandahålla verktyg som abstraherar specifika egenskaper hos arkitekturen. Denna modell är redan etablerad inom grafikområdet, där man kan portera programvara från datorn till olika inbyggnadsplattformar utan att den behöver modifieras. Freescale planerar därför att stödja parallellprogrammeringsspråket OpenCL i sin nästa generation av kameraplattformar. Utmaningen är att bevara hög prestanda och energieffektivitet hos sys-

temkretsen och att stödja standardiserade applikationsgränssnitt (API).

Fram till nu har parallellprogrammeringsspråk som OpenCL inte specificerats med ADAS-system i åtanke. Istället har modellen varit att använda externa minnesbuffer för mellanliggande resultat. Den ökade DRAM-trafiken hos en sådan lösning är ett stort problem i inbyggdsvärlden eftersom den begränsar bandbredden, vilket leder till klart högre effektförbrukning. För att förbättra situationen kan det komma att krävas att dagens standarder utökas för att bättre stödja strömmade databearbetning på chipet.

Genom att utföra flera bildbehandlingssteg utan att belasta externa minnen med delresultat kan man skapa betydligt mer energi- och kostnads-effektiva lösningar. För detta behövs en grafikorienterad bearbetning med kraftigt optimerat dataflöde som är lämpat för de vanliga stegen i bildbehandlingen.

Ytterligare en fördel med att använda ett API från datorvärlden är att man i utvecklingsarbetet kan använda en vanlig pc med ett grafikkort som stöder OpenCL. Det betyder att man kan börja utveckla programvaran långt innan målplattformen är tillgänglig. ■

Detta är SER

SER är föreningen för Sveriges elektro-, data- och IT-ingenjörer.

Vår mission är att stimulera samhällsnyttig utveckling och svenskt näringsliv samt främja den internationella konkurrenskraften för svenska elektro-, data- och IT-ingenjörer!

Mera information om SER finner du på www.ser.se

Eller mejla ser@ser.se!



För smart och hållbar samhällsutveckling

