Relationships Between Computational System Performance and Recognition System Performance

Michael D. DeVore
Joseph A. O'Sullivan
{mdd2,jao}@essrl.wustl.edu
Electronic Systems and Signals Research Laboratory

Roger D. Chamberlain
Mark A. Franklin
{roger,jmf}@ccrc.wustl.edu
Computer and Communications Research Center

Washington University in St. Louis
School of Engineering and Applied Science

Supported by: DARPA grant DAAL01-98-C-0074, ONR grant N00014-98-1-06-06, Boeing Foundation, ARO CIS DAAH04-95-1-0494

Motivation

- ATR systems explicitly or implicitly based on target models with complexity C
- More complex models can yield better accuracy; $\Pr(\text{error}) = f(C, \alpha_{SAR})$
- Complexity and computation power determine throughput; $T_{CHIP} = h(C, \alpha_{COMP})$

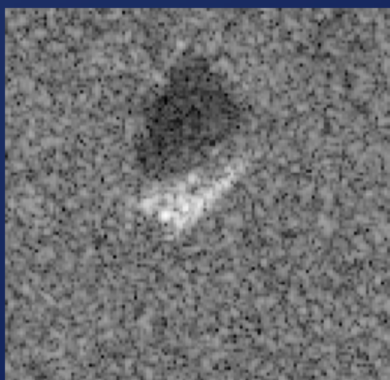
Given an architecture, both accuracy, Pr(error), and throughput, $R=1/T_{CHIP}$, are parameterized by target model complexity

Outline

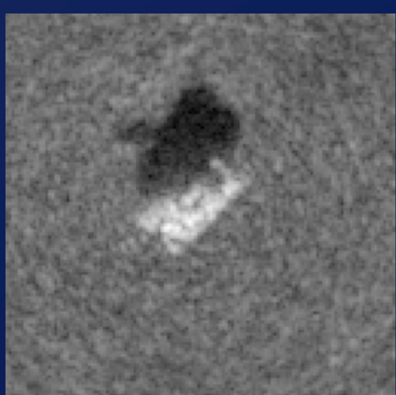
1. Conditionally Gaussian model for SAR
2. Cost of Likelihood Evaluation
3. Approximations and model complexity
4. Computational models
5. Results

Conditionally Gaussian Model

T-72 Photo



SAR Image



Variance Image

Model pixel *i* as independent, zero mean, complex conditionally Gaussian

$$p_{\mathbf{R}|\Theta,A,C^2}\left(\mathbf{r}\,\middle|\,\theta,a,c^2\right)=\prod_i \frac{1}{\pi c^2\sigma_i^2(\theta,a)}\,e^{-\frac{|r_i|^2}{c^2\sigma_i^2(\theta,a)}}$$

Where: a **= target class** $\quad$ σ_i^2 **= variance function**

θ **= target pose** $\quad$ c^2 **= constant (radar power)**

Recognition by maximizing the log-likelihood ratio

$$\left[\hat{a},\hat{\theta},\hat{c}^2\right]=\underset{[a,\theta,c^2]}{\operatorname{argmax}}\ln\left[\prod_i \frac{p\left(\mathbf{r}\,|\,a,\theta,c^2\right)}{p\left(\mathbf{r}\,|\,\xi^2\right)}I_i(a,\theta)\right]$$

Where: ξ^2 **= average clutter variance**

I_i **= mask function**

Likelihood Cost

Let $N(a,\theta)$ = # nonzero mask elements, (i) = ith non-zero mask pixel

$$\left[\hat{a},\hat{\theta}\right] = \underset{[a,\theta]}{\operatorname{argmax}} \left\{ -N(a,\theta)\ln\hat{c}^2 - \sum_{i=1}^{N(a,\theta)} \left[\ln\sigma_{(i)}^2(a,\theta) + 1 - \ln\xi^2\right] + \frac{1}{\xi^2}\sum_{i=1}^{N(a,\theta)} \left|r_{(i)}\right|^2 \right\}$$

$$\hat{c}^2 = \frac{1}{N(a,\theta)} \sum_{i=1}^{N(a,\theta)} \frac{\left|r_{(i)}\right|^2}{\sigma_{(i)}^2(a,\theta)}$$

Expression evaluation consumes (on average):

$2N(a,\theta)$ +2 memory reads totaling $(2N(a,\theta)+2)\,\dfrac{T_{\text{mem}}}{P}$ seconds

$2N(a,\theta)$ +15 operations totaling $(2N(a,\theta)+15)\,\dfrac{\text{CPI}\bullet T_{\text{cyc}}}{P}$ seconds

where: T_{mem} = effective memory read time T_{cyc} = clock period
CPI = average cycles per instruction P = # processors

Successively-Refinable Encoding

Likelihood is a function of $\sigma_i^2(a,\theta)$, a function of continuous θ.
Discrete approximation with N_d non-overlapping intervals of width d:

$$\tilde{\sigma}^2_{d,i}(\theta_k, a) = \frac{1}{d} \int_{\frac{2\pi k}{N_d} - \frac{d}{2}}^{\frac{2\pi k}{N_d} + \frac{d}{2}} \sigma_i^2(\theta, a) d\theta$$

Smaller d yields more complex, piecewise constant representations

Approximations d and $d/2$ are hierarchically related:

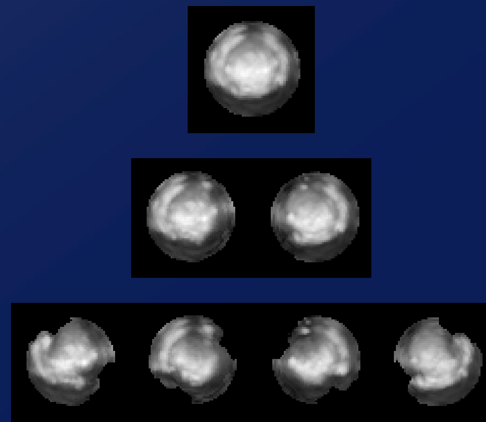
$$\tilde{\sigma}^2_{d,i}(\theta_k, a) = \frac{1}{n_1 + n_2} \left[n_1 \tilde{\sigma}^2_{\frac{d}{2},i}(\theta_{2k}, a) + n_2 \tilde{\sigma}^2_{\frac{d}{2},i}(\theta_{2k+1}, a) \right]$$

Given parent and 1 child, compute other child costing $4N(a,\theta)$ operations saving $b(a,\theta)$/BW in communication time

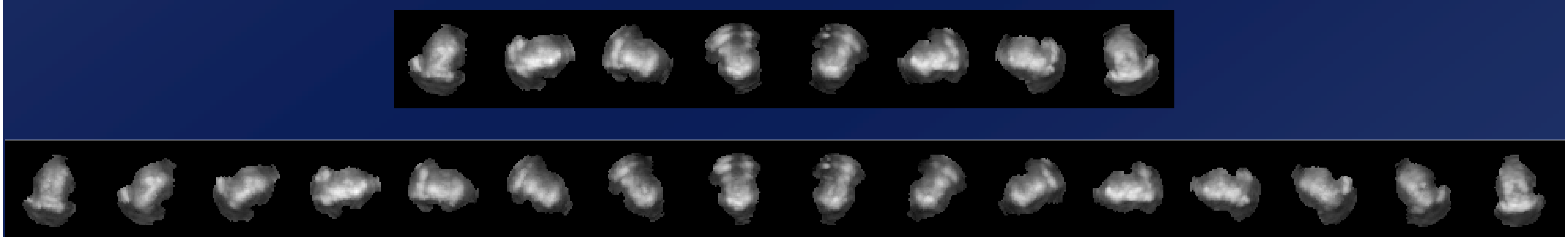
where: $b(a,\theta)$ = bits to represent $\sigma^2(a,\theta)$ and BW = bandwidth

Successively-Refinable Encoding

Consider decreasing interval widths $d_1=2\pi$, $d_2=\pi$, …, $d_m=2\pi/2^{m-1}$



Approximate variance images for bulldozer (D7) from d_1 through d_5



⋮

Search over θ_k in level i ordered by the most likely pose at level i-1

Resource consumption depends on extent of search (complexity)

Computational Model

Time to process through approximation d_m includes time to:

- distribute SAR image to each processor
- process each approximation until local memory is full
- process each remaining approximation

$$T_{\text{chip}} = \frac{S_c}{\text{BW}} \left\lceil \log_2 (P+1) \right\rceil + \sum_{l=1}^{l_{\text{mem}}} 2^{l-1} N_{\text{T}} \tau_{d_l} + \sum_{l=l_{\text{mem}}+1}^{m} 2^{l-2} N_{\text{T}} \left(\tau_{d_l} + \tau'_{d_l} \right)$$

Where:

S_c = bits per SAR image

P = number of processors

BW = network bandwidth

N_T = number of target classes

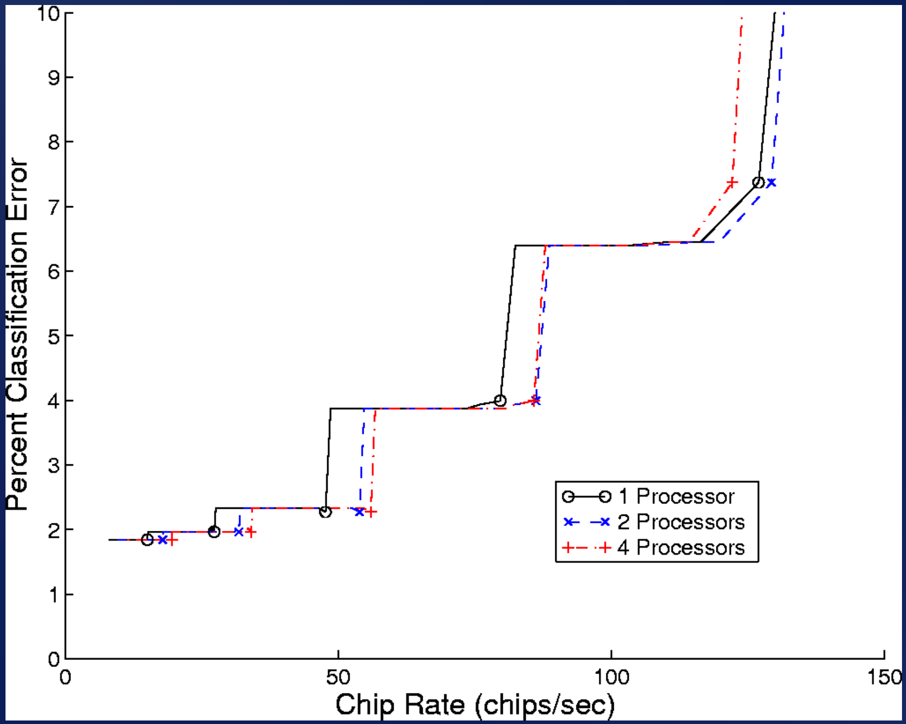
τ_d = average time per template at approximation d exploiting hierarchy. Receive variance, compute variance, and compute likelihoods.

τ'_d = average time per template without exploiting hierarchy. Receive variance and compute likelihood.

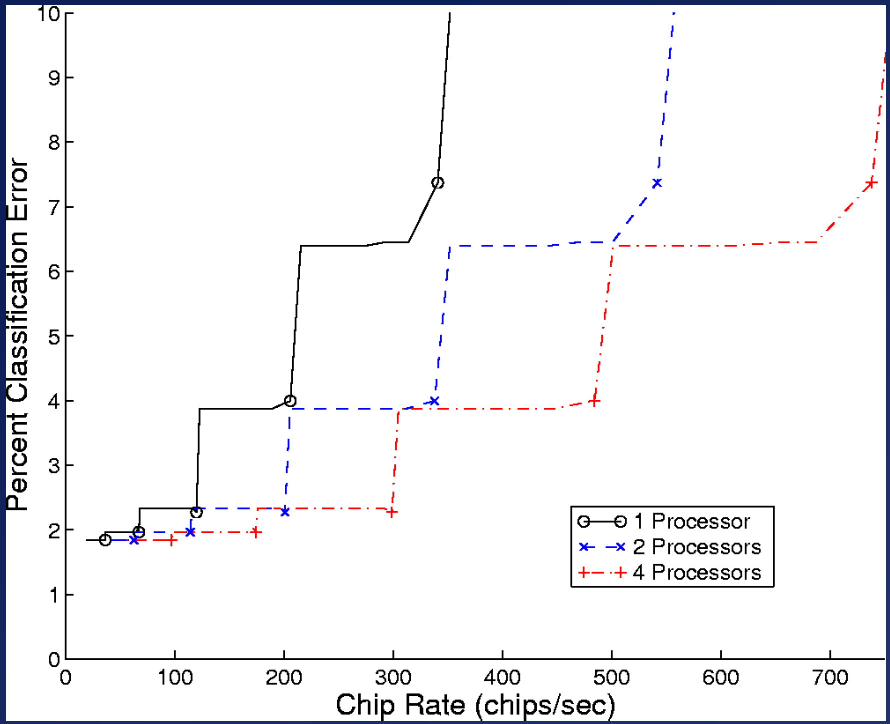
Example

With: Chip Rate = $1/T_{chip}$ 1 GHz clock 64 bit read per clock cycle
4 target classes 0.5 CPI 25 target locations considered

- High throughput corresponds to coarse approximations
- Markers denote start of a new level in model representation tree



1 Gbps interconnection



10 Gbps interconnection

Conclusions

- Process of relating recognition accuracy to throughput
- Unites a computational model with model-based recognition
- Dynamically adjustable operating point
- Accuracy-throughput curves support exploration of design space

Future Work

Variable pixel quantization (bit depth)
Models for multiple, simultaneous recognition objectives
Generic measure of "work" in recognition problems

Appendix

Data Sets

Model	Training Set #1			
	Vehicle	Serial No.	Depression	Images
BMP-2	#1	9563	17°	233
	#2	9566		231
	#3	c21		233
BTR-70	#1	c71	17°	233
BRDM-2	#1	E-71	17°	298
T-72	#1	132	17°	232
	#2	812		231
	#3	s7		228

Target	SOC Testing Set			
	Vehicle	Serial No.	Depression	Images
BMP-2	#1	9563	15°	195
	#2	9566		196
	#3	c21		196
BTR-70	#1	c71	15°	196
BRDM-2	#1	E-71	15°	263
T-72	#1	132	15°	196
	#2	812		195
	#3	s7		191

Throughput Details

$$
\begin{aligned}
T_{\mathrm{chip}} \quad = \quad & \frac{S_{\mathrm{C}}}{\mathrm{BW}}\lceil \log_2(P+1)\rceil + N_{\mathrm{T}}\left[\frac{b_1}{\mathrm{BW}} + \frac{2N_1+2}{P}T_{\mathrm{mem}} + \frac{N_{\mathrm{S}}(2N_1+15)}{P}\cdot \mathrm{CPI}\cdot T_{\mathrm{cyc}}\right] + \\
& N_{\mathrm{T}}\left[\frac{b_1}{\mathrm{BW}} + \frac{4N_1+3}{P}T_{\mathrm{mem}} + \frac{N_{\mathrm{S}}(4N_2+30)+4N_2}{P}\cdot \mathrm{CPI}\cdot T_{\mathrm{cyc}}\right] + \\
& \sum_{l=3}^{m}\left\{2^{l-3}N_{\mathrm{T}}\left[\frac{b_l}{\mathrm{BW}} + \frac{4N_{l-1}+3}{P}T_{\mathrm{mem}} + \frac{N_{\mathrm{S}}(4N_l+30)+4N_l}{P}\cdot \mathrm{CPI}\cdot T_{\mathrm{cyc}}\right] + \right. \\
& \left. \quad 2^{l-3}N_{\mathrm{T}}\left[2\frac{b_l}{\mathrm{BW}} + \frac{3N_l+3}{P}T_{\mathrm{mem}} + \frac{N_{\mathrm{S}}(4N_l+30)}{P}\cdot \mathrm{CPI}\cdot T_{\mathrm{cyc}}\right]\right\}.
\end{aligned}
$$

P	the number of processors in the system
T_{cyc}	cycle time for the processor (sec)
T_{mem}	mean memory access time for a 64 bit floating-point value (sec)
S_{C}	size of a SAR image chip (bits)
BW	bandwidth of the interconnection network (bits/sec)
N_m	the average number of pixels in the segmented templates $\tilde{\boldsymbol{\sigma}}^2_{d_m}(a,\theta_k)$
N_{S}	the number of target positions considered
N_{T}	the number of targets considered
b_m	the average number of bits in the representation of the template $\tilde{\boldsymbol{\sigma}}^2_{d_m}(a,\theta_k)$
CPI	mean cycles per instruction executed by each processor, ignoring memory effects

Table 1. Parameters characterizing the ATR systems under consideration.