Embedding Rectangular Grids into Square Grids with Dilation Two

Rami G. Melhem, *Member, IEEE,* and Ghil-Young Hwang, *Member, IEEE*

***Abstract*—In this paper, a new technique, the multiple ripple propagation technique, is presented for mapping an $h \times w$ grid into a $w \times h$ grid such that the dilation cost is 2, i.e., such that any two neighboring nodes in the first grid are mapped onto two nodes in the second grid that are separated by a distance of at most 2. This technique is then used as a basic tool for mapping any rectangular source grid into a square target grid with the dilation two property preserved. The ratio of the number of nodes in the source grid to the number of nodes in the target grid, called the expansion cost, is shown to be always less than 1.2. This is a significant improvement over the previously suggested techniques, where the expansion cost could be bounded by 1.2 only if the dilation cost was allowed to be as high as 18.**

***Index Terms*—Embeddings among grids, mapping onto meshes, minimizing the dilation cost, squaring up rectangular grids, theory of VLSI layout.**

I. Introduction

In this paper, we study the problem of squaring up a rectangular grid, that is, embedding an $h \times w$ rectangular grid into a $k \times k$ square grid, where $k \geq \lceil \sqrt{hw} \rceil$, and $\lceil \; \rceil$ is the ceiling function. The results of this research may be applied to the VLSI design of highly eccentric circuits that, without squaring up, would have to be laid out in a rectangular area with a height/width ratio deviating significantly from unity [7], [10]. They may also be applied to the mapping of toroidal [6] and rectangular problem domains [3], [8] onto mesh-connected architectures [1], [9]. Mapping of rectangular program graphs onto hypercube architectures may also benefit from this research. Specifically, it has been shown [5] that this mapping may be accomplished by embedding the graph into a square graph which is, then, mapped easily to the hypercube.

Two measures may be used to estimate the quality of an embedding. The first measure is the expansion cost E, which is the ratio of the number of nodes in the square target grid to the number of nodes in the source rectangular grid. That is $E = k^2/hw$. The other measure is the dilation cost D, which is a measure of the communication penalty that has to be paid due to the squaring up. More specifically, if a link λ in the source grid connects two neighboring nodes, say (i, j) and $(i, j+1)$, and these two nodes are mapped to the nodes (i', j') and $(i' + c_i, j' + c_j)$ in the target grid, then the dilation of the edge λ after the embedding is defined by $D(\lambda) = |c_i| + |c_j|$. The dilation cost of the embedding is then given by $D = \max_\lambda D(\lambda)$.

The best known results for embedding an $h \times w$ grid into the smallest possible $k \times k$ grid are given in [2], where different embedding methods are suggested for different ranges of the eccentricity ratio $\rho = w/h$. Assuming that $h \geq 25$, all the methods suggested in [2] produce embeddings with expansion costs smaller than 1.2, and dilation costs ranging from 2 to 18, depending on the value of ρ. Specifically, the dilation cost is less than or equal to 3 if ρ is in one of the ranges $(1, 2]$, $(10/3, 4]$, $(8, 9]$, or $(155, \infty)$. Otherwise, the dilation cost is larger than 5.

In this paper, we first introduce, in Section II, the multiple ripple propagation technique which may be used to embed an $h \times w$ grid onto a $w \times h$ grid with expansion cost 1 and dilation cost 2. This basic technique is then used in Sections III and IV to embed any rectangular grid with $\rho \leq 4$ into a square grid. The idea is to apply the ripple propagation technique to carefully chosen subrectangles of the rectangular grid. For grids with $\rho > 4$, the ripple propagation technique may be combined with the technique of folding [2]. This is described and analyzed in Section V. Finally, in Section VI, we summarize our results and show that it is always possible to square up any rectangular grid at a dilation cost of 2 and an expansion cost less than 1.2. This is a clear improvement over the results given in [2].

II. A Multiple Ripple Propagation Technique

The purpose of the technique described in this section is to map an $h \times w$ grid satisfying

$$h < w \leq 2h \tag{1}$$

onto a $w \times h$ grid with unity expansion cost and with dilation cost equal to 2. In order to accomplish that, the w nodes in each row in the original grid should be compressed to occupy only h columns. For this, we let $l = w - h$, and compress $2l$ nodes from each row into l columns by repeated rippling. The remaining $s = w - 2l = 2h - w$ nodes are left uncompressed. In Fig. 1(b), we show the grid of Fig. 1(a) after compressing each of its rows. As shown in the figure, the positions of the l ripples in each row are chosen as follows. In the first row, the l ripples are grouped to the right, and in the last row, the l ripples are grouped to the left. At each row, one of the ripples, that was grouped to the right in the previous row, starts its propagation to the left (moves one column). The propagation of that ripple continues at a rate of one column every row until it can no longer propagate. The propagation of the ripples is very similar to the motion of the legs of a walking worm.

Fig. 1(b) is laid out to occupy $w + s$ rows and h columns. However, it may be noticed that s positions in each column are not utilized. This allows for the compression of Fig. 1(b) into an $w \times h$ grid which has a dilation cost equal to 2 [see Fig. 1(c)]. In order to be more formal, we let $F(i, j) = (u(i, j), v(i, j))$ be the function which maps each point (i, j) in the source grid to a corresponding point $(u(i, j), v(i, j))$ in the target grid. Here, $(1, 1)$ is the node at the top left corner of the grid. For any node $(1, j)$ in the first row of the source grid, the mapping function F

Manuscript received April 2, 1988; revised June 9, 1989. This work was supported in part by ONR Contract N00014-85-K-0339. A shorter version of this paper appeared in the Proceedings of the Allerton Conference on Computer, Control, and Communications, September 1988.

R. G. Melhem is with the Department of Computer Science, University of Pittsburgh, Pittsburgh, PA 15260.

G.-Y. Hwang is with the Sukuɔ Computer Company, Seoul, Korea.

IEEE Log Number 9038763.

0018-9340/90/1200-1446$01.00 ©1990 IEEE

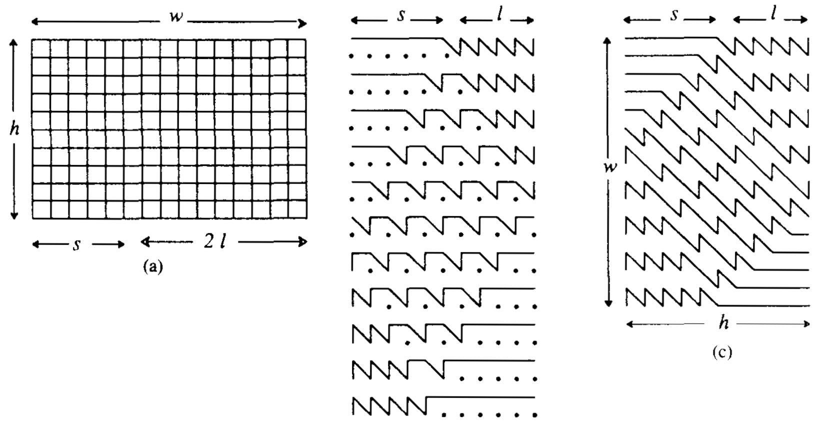
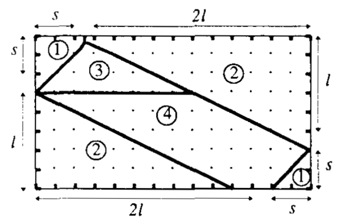
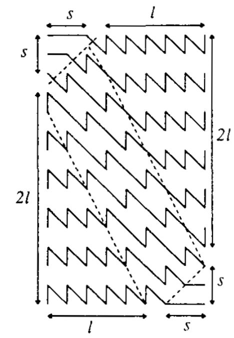


Fig. 1. Embedding an 11 × 16 grid into an 16 × 11 grid using multiple ripple propagation.





Region 1: $j<s-i+2$ OR $j>w-i+l+1$
Region 2: $j>s+2i-2$ OR $j<2i-2s-1$
Region 3: $j\geq s-i+2$ AND $j\leq s+2i-2$ AND $i\leq s+1$
Region 4: $j\geq 2i-2s-1$ AND $j\leq s+2i-2$ AND $i>s+1$ AND $j\leq w-i+l+1$.

Fig. 2. Partitioning of the source grid for $s \leq l$.

is defined as follows:

$$u(1,j)=\begin{cases}1 & \text{if } j=1,\cdots,s\\ 1+\operatorname{rem}\left(\dfrac{j-s}{2}\right) & \text{if } j=s+1,\cdots,w\end{cases} \tag{2.a}$$

$$v(1,j)=\begin{cases}j & \text{if } j=1,\cdots,s\\ s+\left\lceil\dfrac{j-s}{2}\right\rceil & \text{if } j=s+1,\cdots,w.\end{cases} \tag{2.b}$$

where rem( ) denotes the remainder of integer division. The function F may then be defined recursively such that, for any node (i, j) not in the first row, $F(i, j)$ is specified in terms of $F(i-1, j)$. In order to simplify the recursive definition of F, we partition the source grid into four regions as shown in Figs. 2 and 3, and we use different recursive formulas for different regions. Specifically,

$$u(i,j)=\begin{cases}u(i-1,j)+1 & \text{if } (i,j)\in\text{Region 1}\\ u(i-1,j)+2 & \text{if } (i,j)\in\text{Region 2}\\ u(i-1,j)+\Delta_{u,3}(i,j) & \text{if } (i,j)\in\text{Region 3}\\ u(i-1,j)+\Delta_{u,4}(i,j) & \text{if } (i,j)\in\text{Region 4}\end{cases} \tag{3.a}$$

$$v(i,j)=\begin{cases}v(i-1,j) & \text{if } (i,j)\in\text{Region 1}\\ v(i-1,j) & \text{if } (i,j)\in\text{Region 2}\\ v(i-1,j)-\Delta_{v,3}(i,j) & \text{if } (i,j)\in\text{Region 3}\\ v(i-1,j)-\Delta_{v,4}(i,j) & \text{if } (i,j)\in\text{Region 4}\end{cases} \tag{3.b}$$

where $\Delta_{u,3}$ and $\Delta_{v,3}$ depend on the remainder $r(i, j) = \operatorname{rem}((j - s + i - 1)/3)$. Specifically,

$$\Delta_{u,3}=\begin{cases}0 & \text{if } r(i,j)=2\\ 2 & \text{otherwise}\end{cases}$$

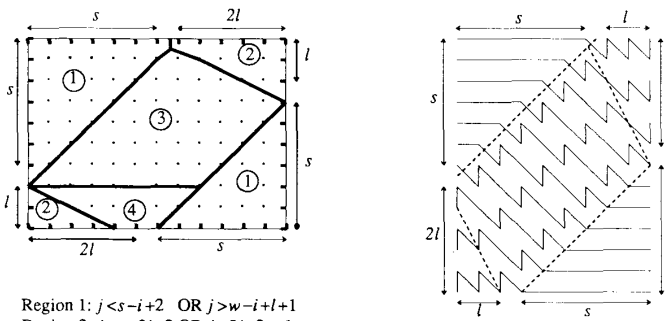
$$\Delta_{v,3}=\begin{cases}1 & \text{if } r(i,j)=2\\ 0 & \text{otherwise.}\end{cases}$$

Similarly, if $\bar{r}(i, j) = \operatorname{rem}((j - 2i + 2s + 2)/3)$, then

$$\Delta_{u,4}=\begin{cases}0 & \text{if } \bar{r}(i,j)=2\\ 2 & \text{otherwise}\end{cases}$$

$$\Delta_{v,4}=\begin{cases}1 & \text{if } \bar{r}(i,j)=2\\ 0 & \text{otherwise.}\end{cases}$$

Given the above formulas, the following theorem proves that the dilation cost of the mapping F is at most two.



Region 1: $j < s-i+2$ OR $j > w-i+l+1$
Region 2: $j > s+2i-2$ OR $j < 2i-2s-1$
Region 3: $j \geq s-i+2$ AND $j \leq s+2i-2$ AND $i \leq s+1$ AND $j \leq w-i+l+1$
Region 4: $j \geq 2i-2s-1$ AND $j \leq w-i+l+1$ AND $i > s+1$.

Fig. 3. Partitioning of the source grid for $s > l$.

Theorem: For any (i, j), where $i, j > 1$, the following is true

$$|u(i, j) - u(i-1, j)| + |v(i, j) - v(i-1, j)| \leq 2 \quad (4.a)$$

and

$$|u(i, j) - u(i, j-1)| + |v(i, j) - v(i, j-1)| \leq 2. \quad (4.b)$$

That is, any two adjacent nodes in the source grid are mapped onto two nodes whose separation distance in the target grid is less than or equal to 2.

Proof: The proof of (4.a) is straightforward if (i, j) is in Region 1 or Region 2. If (i, j) is in region 3, then the left side of (4.a) reduces to $|\Delta_{u,3}(i, j)| + |\Delta_{v,3}(i, j)|$, which is equal to 1 if $r(i, j) = 2$, and to 2 otherwise. The case $(i, j) \in$ Region 4 is similar.

To prove (4.b) we use induction on i. For $i = 1$, the proof is by direct substitution from (2). Next, assuming that (4.b) holds for $i - 1$, we should show that it also holds for i. Again the inductive proof is straightforward if (i, j) is in Regions 1 or 2, and is similar if (i, j) is in Regions 3 and 4. For this reason we will consider in the rest of this proof only the case in which $(i, j) \in$ Region 3. For this case, we will prove, by induction, a more restrictive form of (4.b), namely

$$u(i, j) - u(i, j-1) = \begin{cases} -1 & \text{if } r(i, j) = 2 \\ 1 & \text{if } r(i, j) = 0 \text{ or } 1 \end{cases} \quad (5.a)$$

$$v(i, j) - v(i, j-1) = \begin{cases} 0 & \text{if } r(i, j) = 2 \\ 1 & \text{if } r(i, j) = 0 \text{ or } 1. \end{cases} \quad (5.b)$$

For $i = 2$, (5) is proved directly from (2) and (3). To prove the induction step, we notice that if (i, j) is in Region 3, then $(i, j-1)$ is either in Region 3 or in Region 1. We first assume that $(i, j-1)$ is in Region 3 and use (3) to obtain

$$u(i, j) - u(i, j-1) = u(i-1, j) - u(i-1, j-1) + \Delta_{u,3}(i, j) - \Delta_{u,3}(i, j-1) \quad (6.a)$$

$$v(i, j) - v(i, j-1) = v(i-1, j) - v(i-1, j-1) - \Delta_{v,3}(i, j) + \Delta_{v,3}(i, j-1). \quad (6.b)$$

If $r(i, j) = 2$, then $\Delta_{u,3}(i, j) = 0$ and $\Delta_{u,3}(i, j-1) = 2$ because $r(i, j-1) = 1$. Also, $r(i-1, j) = 1$, which, from the induction hypothesis, gives $u(i-1, j) - u(i-1, j-1) = 1$. Therefore, from (6.a) we have $u(i, j) - u(i, j-1) = 1 + 0 - 2 = -1$. Similarly, $\Delta_{v,3}(i, j) = 1$, $\Delta_{v,3}(i, j-1) = 0$, and $v(i-1, j) - u(i-1, j-1) = 1$, from which we obtain $v(i, j) - v(i, j-1) = 1 - 1 + 0 = 0$. A similar argument applies if $r(i, j) = 0$ or 1.

Finally, if $(i, j-1)$ is in Region 1, then $j = s - i + 2$ and thus, $r(i, j) = 1$. From (3.a), we get $u(i, j) - u(i, j-1) = u(i-1, j) - u(i-1, j-1) + 2 - 1$ and from (3.b) we get $v(i, j) - v(i, j-1) = v(i-1, j) - v(i-1, j-1)$. But both $(i-1, j)$ and $(i-1, j-1)$ are in Region 1, and thus $u(i-1, j) = u(i-1, j-1)$, and $v(i-1, j) = v(i-1, j-1) + 1$, which proves (5.a) and (5.b), respectively. □

The above theorem proves that it is possible to map an $h \times w$ grid, $h < w \leq 2h$, exactly into a $w \times h$ grid with dilation cost 2. It is also possible to concatenate the $w \times h$ target grid with its symmetric image (reflected across the line $v = h$) to obtain an exact embedding of an $h \times 2w$ grid into a $w \times 2h$ grid with dilation cost 2. Along the same line of thinking, an $h \times 2w + 1$ source grid may be divided into $h \times w + 1$ and $h \times w$ subgrids. These two subgrids may then be embedded into a $w + 1 \times h$ and a $w \times h$ grid, respectively, and by concatenating the former with the symmetric image of the latter, we may obtain a $w + 1 \times 2h$ target grid. The dilation cost at the line of concatenation may be shown to be at most two. Grid concatenations of the type described here will be used repeatedly and tacitly in the rest of this paper.

In the following sections, we apply the above technique to our original problem of mapping an $h \times \rho h$ rectangular grid (ρ is assumed to be greater than unity), into a square grid. First, two basic methods are introduced for grids with $\rho \leq 4$. These methods are then combined with folding and applied effectively to the embedding of any grid with $\rho > 4$.

III. The Method of Exact Row Fitting

Let k, $k \geq h$, be the dimension of the square grid (called the target grid) onto which a given $h \times \rho h$ grid (called the source grid) is to be mapped. Of course, it is desirable to choose the smallest possible k in order to minimize the expansion cost $E = k^2/\rho h^2$. Given such a k, the method of exact row fitting assumes that the leftmost $h \times k$ subgrid of the source grid may be mapped exactly into the $k \times h$ leftmost subgrid of the target grid (see Fig. 4). This is possible if and only if condition (1) is

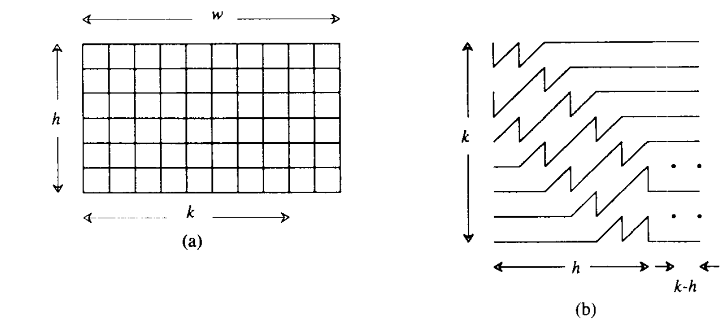


Fig. 4. Embedding a 7 × 11 grid into a 9 × 9 grid using the method of exact row fitting.

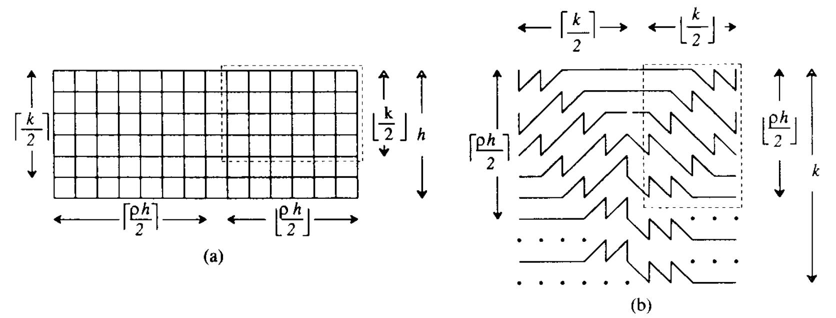


Fig. 5. Embedding a 7 × 15 grid into a 11 × 11 grid using exact column fitting.

satisfied. That is

$$h \le k \le 2h. \tag{7.a}$$

Moreover, if

$$k - h \ge \rho h - k \tag{7.b}$$

that is, the number of columns, $k - h$, remaining in the target grid is at least equal to the number of columns, $\rho h - k$, remaining in the source grid, then these columns may be mapped in a trivial way with dilation cost 2. In other words, the mapping may be completed with dilation cost 2 provided that the size of the target grid, k, satisfies the conditions (7.a/b).

The solution of inequalities (7.a/b) may be found by, first, computing the minimum k that satisfies (7.b), and then checking that this value is consistent with (7.a). Specifically, (7.b) is satisfied if

$$k = K_r = \left\lceil \frac{\rho + 1}{2} h \right\rceil. \tag{8}$$

It is straightforward to check that the value of k given by (8) satisfies (7.a) if $\rho \le 3 - 1/h$. Hence, the method of exact row fitting may be applied only if $\rho \le 3 - 1/h$. Noting that $K_r \le ((\rho + 1)h + 1)/2$, we may obtain an upper bound on the expansion cost of the resulting embedding. Namely,

$$E_r = \frac{K_r^2}{\rho} \le E_{r,\max} = \frac{\left(\rho + 1 + \frac{1}{h}\right)^2}{4\rho}. \tag{9}$$

The value of E_r increases monotonically with ρ for $\rho > 1 + 1/h$, and hence may exceed 1.2 for large values of ρ. For examples, assuming $h = 12$, then $E_r > 1.2$ if $\rho \ge 2.06$. Moreover, if $\rho > 3 - 1/h$, the method may not be applied. In these cases, the method of exact column fitting, described in the following section, can be used.

IV. The Method of Exact Column Fitting

The embedding technique used in this section is based on the vertical dissection of both the source and the target grids, each into two subgrids which are as equal as possible. Each of the source subgrids is then embedded into the corresponding target subgrid in a way that ensures that all the columns of the target grid are efficiently used. In order to deal with the case of ρh being an odd integer, the number of columns in the two source subgrids is taken to be $\lceil \rho h/2 \rceil$ and $\lfloor \rho h/2 \rfloor$, respectively, where $\lfloor\ \rfloor$ is the floor function. For the same reason, the number of columns in the target subgrid is divided into $\lceil k/2 \rceil$ and $\lfloor k/2 \rfloor$ columns, respectively, (see Fig. 5).

The optimal size, $k = K_c$, of the target grid should be determined by the embedding of $h \times \lceil \rho h/2 \rceil \to k \times \lceil k/2 \rceil$ or the embedding $h \times \lfloor \rho h/2 \rfloor \to k \times \lfloor k/2 \rfloor$, whichever gives a more strict condition on k. It turns out that the latter embedding is more restrictive than the former, and hence, should be used to derive k. In the remainder of this section, we will denote the $h \times \lfloor \rho h/2 \rfloor$ grid by G_s, and the $k \times \lfloor k/2 \rfloor$ grid by G_t, and we will describe an embedding of G_s into G_t. The embedding of the other half of the source grid (the $h \times \lceil \rho h/2 \rceil$ subgrid) into the other half of the target grid (the $k \times \lceil k/2 \rceil$ subgrid) may be accomplished in a similar fashion.

Consider the upper $\lfloor k/2 \rfloor \times \lfloor h/2 \rfloor$ subgrid of G_s, and embed it into the upper $\lfloor \rho h/2 \rfloor \times \lfloor k/2 \rfloor$ subgrid of G_t using the ripple propagation technique of Section II. In order to

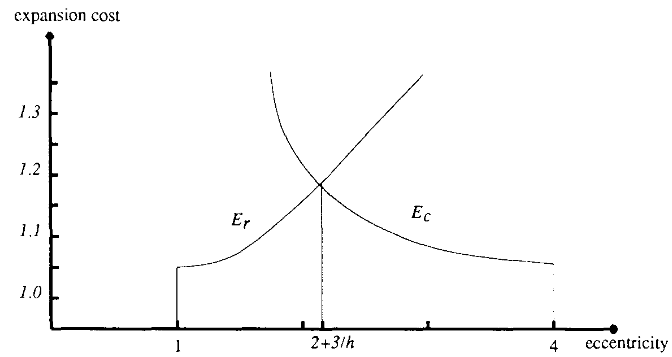


Fig. 6. Expansion cost versus eccentricity for $1 < \rho \leq 4$ ($h = 20$).

accomplish this embedding we should have

$$\lfloor k/2 \rfloor \leq h \tag{10.a}$$

and condition (1) should be satisfied, namely

$$\lfloor k/2 \rfloor \leq \lfloor \rho h/2 \rfloor \leq 2 \lfloor k/2 \rfloor . \tag{10.b}$$

With this, each of the remaining $h - \lfloor k/2 \rfloor$ rows in G_s may then be compressed to have the same pattern as the last row of the $\lfloor \rho h/2 \rfloor \times \lfloor k/2 \rfloor \to \lfloor k/2 \rfloor \times \lfloor \rho h/2 \rfloor$ embedding. This results in a dilation cost equal to 2 and requires $2(h - \lfloor k/2 \rfloor)$ additional rows in G_t. Thus, the following should be satisfied:

$$k - \lfloor \rho h/2 \rfloor \leq 2(h - \lfloor k/2 \rfloor). \tag{10.c}$$

Noting that $(x+1)/2 \geq \lfloor x/2 \rfloor \geq (x-1)/2$, we may calculate the minimum value of k which always satisfies (10.c). Namely,

$$k = K_c = \left\lceil \frac{\rho h/2 + 2h + 1}{2} \right\rceil. \tag{11}$$

This value of K_c satisfies conditions (10.a) and (10.b) as long as $\rho \leq 4$.

With the value of k given by (11), the two halves of the source grid may be successfully embedded into the two halves of the target grid with dilation cost 2. Noting that $\lceil x/4 \rceil \leq (x + 3)/4$, it is possible to bound the expansion cost of the embedding as follows:

$$E_c \leq E_{c,\max} = \frac{\left(\rho + 4 + \frac{5}{h}\right)^2}{16\rho}. \tag{12}$$

For $1 < \rho \leq 4$, the value of E_c is monotonically decreasing with ρ, which suggests the use of the method of fitting columns whenever the method of fitting rows fails to satisfy $E_r \leq 1.2$ (see Fig. 6). The critical value of ρ that determines which of the two methods has a smaller expansion cost may be found by solving $E_{r,\max} = E_{c,\max}$. From (9) and (12), this gives $\rho = 2 + 3/h$. The expansion cost at this value of ρ is $(3h + 4)^2/(4h(2h+3))$, which is always smaller than 1.2 if $h > 17$.

Hence, for $1 < \rho \leq 4$, the most efficient embedding method depends on the value of ρ. Specifically, if $\rho \leq 2 + 3/h$, then the methods of exact row fitting should be used, otherwise, the method of exact column fitting should be used. For values of ρ larger than four, the above methods can be combined with the known method of folding [2] as described in the next section.

V. Combining Ripple Propagation with Folding

If $\rho = (q+1)^2$ for some integer $q \geq 1$, then the source grid may be folded $q + 1$ times to fit exactly an $(q+1)h \times (q+1)h$ target grid. In fact, it is easy to show that if

$$\frac{(q+1)^2}{1.2} < \rho \leq (q+1)^2 \tag{13.a}$$

for some integer $q \geq 1$, then folding the source grid into an $(q+1)h \times (q+1)h$ target grid will result in an expansion cost less than 1.2. In Fig. 7, we illustrate the technique of folding by an example. As clear from this figure, successive tracks (a track consists of h consecutive rows of the target grid) are joined by two $h \times h$ corner tiles that guarantee a dilation cost equal to two.

As described above, folding may result in few unused columns in the last track of the target grid, and condition (13.a) limits the number of these unused columns. It is also possible to apply folding and leave some rows of the target grid unused. More precisely, if the eccentricity of the source grid satisfies

$$q^2 \leq \rho \leq 1.2q^2 \tag{13.b}$$

then it is possible to fold this grid into an $\rho h/q \times \rho h/q$ target grid. This will leave $(\rho - q^2)h/q$ unused rows in the target grid, and condition (13.b) will guarantee that the number of unused rows does not exceed $0.2qh$. Thus, the expansion cost will be less than 1.2.

A. Combining Folding with Exact Row Fitting

Consider an $h \times \rho h$ source grid which satisfies $1.2q^2 < \rho < (q+1)^2/1.2$. Clearly, folding this grid into a square grid is too expensive (expansion cost larger than 1.2) because neither (13.a) nor (13.b) is satisfied. In this section, we introduce a method which combines folding and exact row fitting. This method will be denoted by FR. In order to describe the FR method we assume that the source grid is to be embedded into a target grid of size k, where k satisfies $qh \leq k \leq (q+1)h$. The embedding starts by folding the source grid into the target grid q times as shown in Fig. 8(a). Clearly, the rightmost $h \times (\rho h - qk)$ subgrid of the source grid will not fit into the target grid, and the last $k - qh$ rows of the target grid will be unused. The idea is to consider the last track resulting from the folding (an $h \times \rho h - (q-1)k$ grid denoted by G_s), and to squeeze it into an $k - (q-1)h \times k$ grid (denoted by G_t) that fits the target grid.

The squeeze is performed by partitioning G_s vertically into P subgrids $G_{s,0}, \cdots, G_{s,P-1}$, and partitioning G_t vertically into

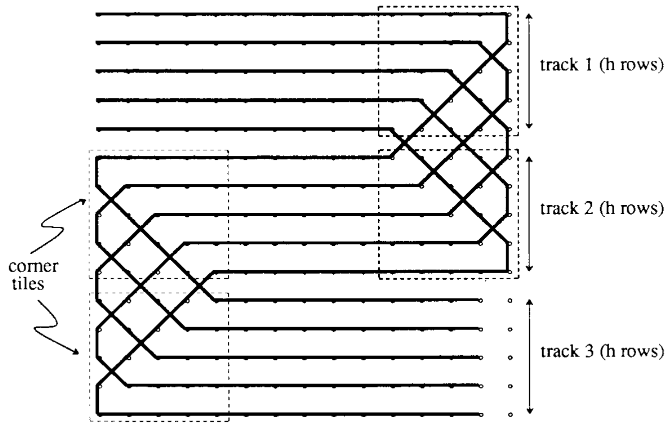


Fig. 7. Folding a 5 × 40 grid into a 15 × 15 grid.

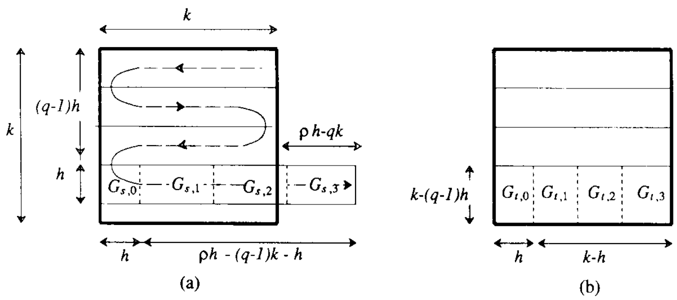


Fig. 8. Combining folding with exact row fitting ($q = 4$ and $P = 4$).

P subgrids, $G_{t,0}, \cdots, G_{t,P-1}$, and then mapping each $G_{s,i}$ into the corresponding $G_{t,i}$. The partitioning of G_s is such that $G_{s,0}$, is an $h \times h$ grid and each remaining $G_{s,i}$, $i = 1, \cdots, P-1$ is an $h \times \lceil (\rho h - (q-1)k - h)/(P-1) \rceil$ grid. Note that if $\rho h - (q-1)k - h$ does not divide $P-1$ then $G_{s,P-1}$ will have few empty columns. Similarly, the partition of G_t is such that $G_{t,0}$ is a $k - (q-1)h \times h$ grid and each of the remaining $G_{t,i}$, $i = 1, \cdots, P-1$ is a $k - (q-1)h \times \lfloor (k-h)/(P-1) \rfloor$ grid.

The method of exact row fitting introduced in Section III is used to map each $G_{s,i}$, $i = 1, \cdots, q-1$, into the corresponding $G_{t,i}$. As for the mapping $G_{s,0} \rightarrow G_{t,0}$, it should ensure that the transition from track $q-1$ to track q does not increase the dilation cost beyond two. This may be accomplished by expanding the $h \times h$ corner tiles (see Fig. 7) into a $k - (q-1)h \times h$ pattern that fits $G_{t,0}$ such that the distribution of the h nodes in the last column of $G_{t,0}$ is similar to the distribution of the h nodes in the first column of $G_{t,1}$. This is always possible when exact row fitting is used to map $G_{s,1}$ to $G_{t,1}$. Specifically, as a result of exact row fitting, the nodes $n_1, \cdots, n_h$ in the first column of $G_{s,1}$ are mapped to the nodes $F(n_1), \cdots, F(n_h)$ in the first column of $G_{t,1}$ such that one of the following two conditions is satisfied for some integer z, $1 \le z \le h$:

1) the distance between $F(n_i)$ and $F(n_{i+1})$ is one, for $i = 1, \cdots, z-1$, and two, for $i = z, \cdots, h-1$.

2) the distance between $F(n_i)$ and $F(n_{i+1})$ is two, for $i = 1, \cdots, h-z$, and one, for $i = h-z+1, \cdots, h-1$.

For example, in the mapping of Fig. 4, $h = 7$, $z = 5$, and the second case applies. In general, using the notation in Fig. 4, the value of z may be found from

$$z = 2h - k.$$

Because of the above property, it is straightforward to expand the $h \times h$ corner tile $G_{s,0}$ into $G_{t,0}$ such that the maximum dilation in $G_{t,0}$ is two and the node distribution in the last column of $G_{t,0}$ is identical to the node distribution in the first column of $G_{t,1}$. In Fig. 9, we show the expansion of a 7×7 corner tile to match the grid of Fig. 4(b).

In order to compute the optimum size k of the target grid, we follow the same reasoning as in Section III. Specifically, the method of exact row fitting may be used for mapping any $G_{s,i}$ into the corresponding $G_{t,i}$. For this, the following conditions should be satisfied (refer to Fig. 10):

$$h \le \left\lfloor \frac{k-h}{P-1} \right\rfloor \tag{14.a}$$

$$h \le k - (q-1)h \le 2h \tag{14.b}$$

$$\left\lfloor \frac{k-h}{P-1} \right\rfloor - h \ge \left\lceil \frac{(\rho - 1)h - (q-1)k}{P-1} \right\rceil - (k - (q-1)h). \tag{14.c}$$

In order to solve the above system of inequalities, we first find the minimum value of k which always satisfies (14.c). This value is

$$k = K_{fr} = \left\lceil \frac{(\rho + q(P-1))h + 2(P-2)}{q+P-1} \right\rceil. \tag{15}$$

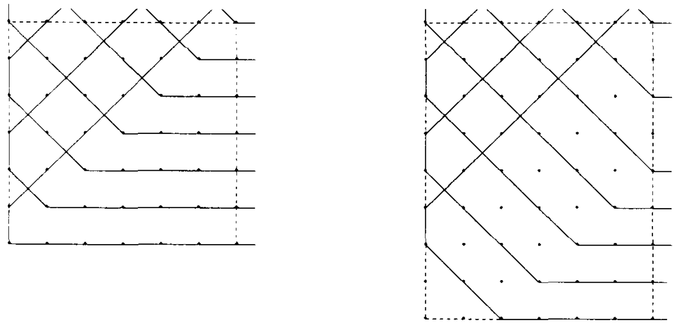


Fig. 9. Mapping a 7 × 7 corner tile into a 9 × 7 grid.

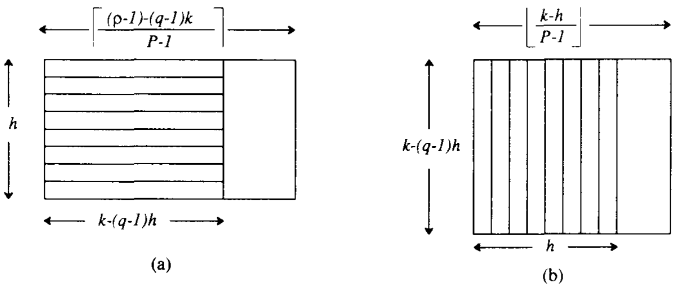


Fig. 10. Mapping $G_{s,i}$ into $G_{t,i}$ in the FR method.

By substituting (15) in (14.a) and (14.b), we conclude that these two conditions are satisfied, respectively, if

$$\rho \geq P^2 + q - P + \frac{(P+q-3)(P-2)}{h} \tag{16.a}$$

and

$$\rho \leq q(q+1) + P - 1 - \frac{3P+q-6}{h}. \tag{16.b}$$

Hence, given any rectangular grid, the method may be used if there exists a P that satisfies (16). The number of partitions P also affects the expansion cost. More precisely, from (15), we find that $K_{fr} \leq (\rho h + q(P-1)h + 3P + q - 6)/(q + P - 1)$, which may be used to bound the expansion cost by

$$E_{fr} \leq E_{fr,\max} = \frac{(\rho + q(P-1) + (3P+q-6)/h)^2}{(P+q-1)^2 \rho}. \tag{17}$$

The derivative $\partial E_{fr,\max}/\partial P$ is negative for $\rho \geq q^2$ which means that, from the point of view of minimizing E_{fr}, it is advantageous to find the maximum P which satisfies (16). For $P > q$, the two conditions (16.a) and (16.b) may not be satisfied simultaneously. If, however, ρ is in the range

$$q^2 + \frac{(q-2)(2q-3)}{h} \leq \rho \leq (q+1)^2 - 2 - \frac{4q-6}{h}, \tag{18}$$

then (16.a/b) are satisfied for $P = q$, and hence the embedding may be completed with q partitions in a target grid whose size is given by (15). The maximum expansion cost may then be found by substituting $P = q$ in (17) to obtain

$$E_{fr} \leq E_{fr,\max} = \frac{(\rho + q(q-1) + (4q-6)/h)^2}{(2q-1)^2 \rho}. \tag{19}$$

In Section VI, it will be shown that, for $q \geq 3$, E_{fr} is smaller than 1.2 for any ρ in the range specified by (18), and that, outside that range, ρ satisfies (13.a) or (13.b), which means that folding may be used with expansion cost less than 1.2. The case $q = 2$, however, is slightly more complicated. For instance, if $h \geq 20$, then, folding is too expensive in the range $5.85 \leq \rho \leq 7.5$ (expansion cost is larger than 1.2). Also, in that range, the FR method either does not apply (if $\rho > 6.9$) or gives $E_{fr} > 1.2$ (if $5.85 < \rho < 6.9$). In this case, combining folding with the method of exact column fitting (the FC method) turns out to be useful. Although we only need this combination for $q = 2$, the FC method will be described in the next section for general q. The reason for doing so is that for $q \geq 3$, although both the FC and the FR methods realize an expansion cost less than 1.2, it will be shown that the FC method gives better results than the FR method for some subranges of ρ.

B. Combining Folding with Exact Column Fitting

In this method, denoted from now on by FC, the source grid is folded into the target grid as described in the previous section, and also each of G_s and G_t is partitioned into P subgrids. The FC method is different from the FR method in that each subgrid $G_{s,i}$, $i = 1, \cdots, P-1$, is mapped into the corresponding $G_{t,i}$ using the method of exact column fitting rather than exact row fitting.

The conditions that have to be satisfied in order to map $G_{s,i}$ into $G_{t,i}$ using exact column fitting are analogous to the conditions (10.a/b/c) of Section IV. Specifically, these conditions are

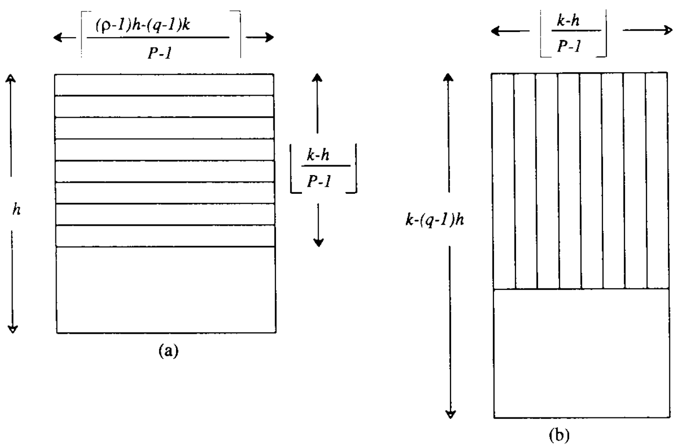


Fig. 11. Mapping $G_{s,i}$ into $G_{t,i}$ in the FC method.

(refer to Fig. 11):

$$\left\lceil \frac{(\rho-1)h-(q-1)k}{P-1} \right\rceil \le k-(q-1)h \quad (20.a)$$

$$\left\lfloor \frac{k-h}{P-1} \right\rfloor \le \left\lceil \frac{(\rho-1)h-(q-1)k}{P-1} \right\rceil \le 2\left\lfloor \frac{k-h}{P-1} \right\rfloor \quad (20.b)$$

$$k-(q-1)h-\left\lceil \frac{(\rho-1)h-(q-1)k}{P-1} \right\rceil \ge 2\left(h-\left\lfloor \frac{k-h}{P-1} \right\rfloor\right). \quad (20.c)$$

The same technique that was used in the last sections is applied to the solution of the above inequalities. From (20.c), the minimum size of the target grid is found to be

$$k = K_{fc} = \left\lceil \frac{(\rho+qP+P-q)h+3(P-2)}{P+q} \right\rceil. \quad (21)$$

By substituting (21) in (20.a) we obtain the condition

$$\rho \le P^2 - P + q + (P+q-3)(P-2)/h$$

which may be satisfied only if $P > q$. Also, by using (21) to compute the expansion cost E_{fc}, and then differentiating the resulting formula, we find that $\partial E_{fc}/\partial P$ is positive for $\rho \le P^2 - P + q$. This means that using $P = q + 1$ partitions will give the best expansion cost. Now using $P = q + 1$ in (21), and substituting the result in (20.a) and (20.b), we find that these conditions are satisfied if ρ lies in the following range

$$q^2 + 1 + \frac{3(q-1)}{h} \le \rho \le (q+1)^2 - 1 - \frac{(q-1)^2}{qh}. \quad (22)$$

That is, the FC method may be applied if ρ satisfies (22). The expansion cost may then be computed from (21) with $P = q + 1$. The upper bound on this cost is given by

$$E_{fc} \le E_{fc,\max} = \frac{\left(\rho + q(q+1) + 1 + (5q-3)/h\right)^2}{(2q+1)^2 \rho}. \quad (23)$$

VI. Discussion and Conclusion

Given an $h \times \rho h$ source grid, let q be the integer that satisfies $q^2 \le \rho < (q+1)^2$. For $q = 1$, it has been shown in Sections III and IV that the mapping of the source grid into a square rectangular grid may be accomplished by using the method of exact row fitting if $\rho \le 2 + 3/h$, or the method of exact column fitting if $\rho > 2 + 3/h$. In both cases, the expansion cost is proven to be less than 1.2.

For $q \ge 2$, the FR or the FC methods described in Section V may be applied provided that $\rho_1 < \rho < \rho_2$, where the critical values ρ_1 and ρ_2 are specified from (18) and (22). Namely,

$$\rho_1 = q^2 + \frac{(q-2)(2q-3)}{h} \quad (24.a)$$

$$\rho_2 = (q+1)^2 - 1 - \frac{(q-1)^2}{qh}. \quad (24.b)$$

In order to determine which of the two methods gives a smaller expansion cost, we notice from (19) and (23) that, for $\rho > 0$, $E_{fr,\max}$ and $E_{fc,\max}$ intersect at only one point, namely

$$\rho_3 = q(q+1) - \frac{1}{2} + \frac{2q^2 - 3q + 9}{2h}. \quad (24.c)$$

We also observe that both $E_{fr,\max}$ and $E_{fc,\max}$ are of the form $f(\rho) = (\rho + a)^2/(b\rho)$, for some constants a and b. Given that, for $\rho > 0$, the function f has only one local minimum at $\rho = a$, we may determine that $E_{fr,\max}$ has its local minimum at $\rho_{fr} = q(q-1) + (4q-6)/h$, and $E_{fc,\max}$ has its minimum at $\rho_{fc} = q(q+1) + 1 + (5q-3)/h$. Clearly, ρ_{fr} is smaller than ρ_1 for $h > 4$, and ρ_{fc} lies between ρ_1 and ρ_2. This leads to the conclusion that $E_{fc,\max} < E_{fr,\max}$ if $\rho > \rho_3$ and $E_{fc,\max} > E_{fr,\max}$ if $\rho < \rho_3$. In Fig. 12, both $E_{fc,\max}$ and $E_{fr,\max}$ are plotted for $q = 2$, $h = 20$, and for $q = 4$, $h = 10$.

Hence, if ρ lies between ρ_1 and ρ_3, the FR method is

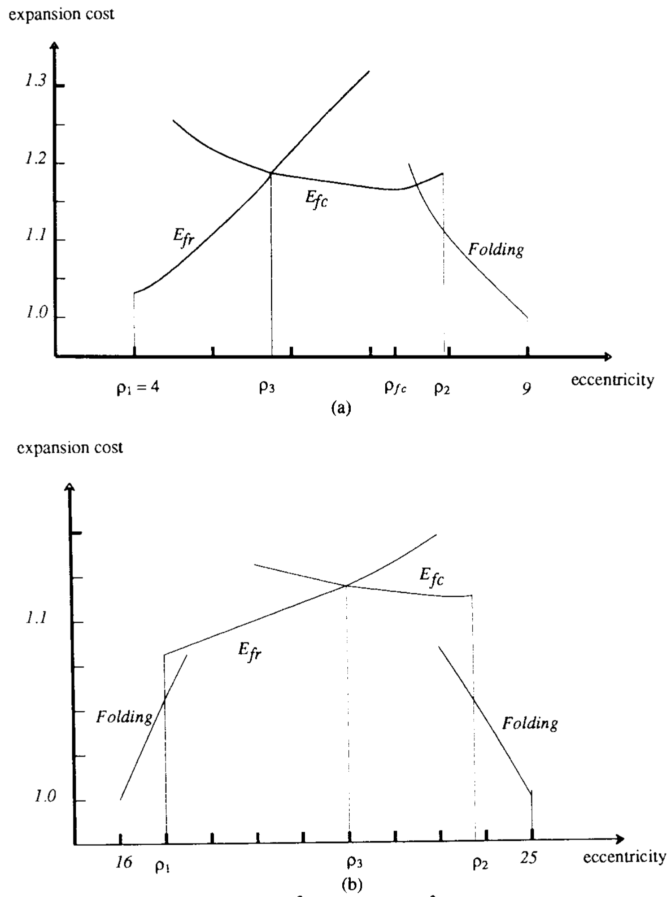


Fig. 12. Expansion cost for $q^2 \leq \rho < (q+1)^2$, (a) $q = 2$, $h = 20$. (b) $q = 4$, $h = 10$.

recommended, and if ρ lies between ρ_3 and ρ_2, then the FC method is recommended. If this strategy is applied, then the largest expansion cost occurs at either $\rho = \rho_2$ or $\rho = \rho_3$. By direct substitution of (24.b) and (24.c) into (23), and after simple algebraic manipulation, it may be shown that, for $h \geq 18$, the value of $E_{fc,\max}$ is less than 1.2 at ρ_2 and ρ_3.

Neither the FR method nor the FC method may be applied if ρ is less than ρ_1 or larger than ρ_2. However, in these two cases, the expansion cost resulting from simple folding is low because ρ is close enough to q^2 and $(q+1)^2$, respectively. In fact, if $\rho \leq \rho_1$, then $\rho < q^2(1 + 2/h)$, which satisfies (13.b) if $h > 10$. Also, if $\rho \geq \rho_2$, then $\rho \geq (q+1)^2(1 - 1/qh) - 1$, which satisfies (13.a) if $h > 10$. In other words, the application of simple folding in these two regions will result in an expansion cost less than 1.2.

In brief, new techniques have been presented and analyzed in this paper, for embedding an $h \times \rho h$ rectangular grid into a square grid with dilation cost equal to two. The most appropriate technique for a given grid has been shown to depend on the size of that grid, that is on h and ρ. By adhering to the selection strategy suggested in the paper, the expansion cost is guaranteed to be smaller than 1.2 if h is larger than or equal to 18.

Finally, it should mentioned that, recently, the same problem of embedding rectangular grids into square grids has been studied independently by Ellis [4]. Specifically, it is shown in [4] that it is possible, with dilation at most three, to embed any $h \times w$ grid into a $k+1 \times k+1$ grid, where $k = \lceil \sqrt{hw} \rceil$.

References

[1] L. Adams and R. Voigt, "Design, development and use of the finite element machine," in *Large Scale Scientific Computation*, S. Parter, Ed. New York: Academic, 1984, pp. 301–321.

[2] R. Aleliunas and A. Rosenberg, "On embedding rectangular grids in square grids," *IEEE Trans. Comput.*, vol. C-31, no. 9, pp. 907–913, 1982.

[3] S. Bokhary, "On the mapping problem," *IEEE Trans. Comput.*, vol. C-30, no. 3, pp. 207–214, 1981.

[4] J. Ellis, "Embedding rectangular grids into square grids," in *Proc. Aegean Workshop Comput.*, June 1988.

[5] S. L. Johnsson, "Communication efficient basic linear algebra computations on hypercube architectures," *J. Parallel Distributed Comput.*, vol. 4, pp. 133–172, 1987.

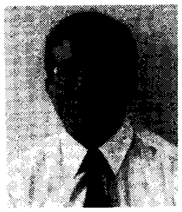
[6] Y. Ma and L. Tao, "Embedding among toruses and meshes," in *Proc. 1987 Int. Conf. Parallel Processing*, pp. 178–187.

[7] C. Mead and L. Conway, *Introduction to VLSI Systems.* Reading, MA: Addison-Wesley, 1980.

[8] P. Sadayappan and F. Ercal, "Nearest-neighbor mapping of finite element graphs onto processor meshes," *IEEE Trans. Comput.*, vol. C-36, no. 12, pp. 1408–1424, 1987.

[9] P. Tseng, K. Hwang, and P. Kumar, "A VLSI based multiprocessor architecture for implementing parallel algorithms," in *Proc. Int. Conf. Parallel Processing*, 1985, pp. 657–664.

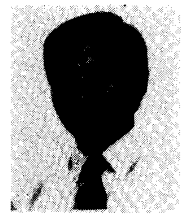
[10] L. Valiant, "Universality considerations in VLSI circuits," *IEEE Trans. Comput.*, vol. C-30, pp. 135-140, 1981.



Rami G. Melhem (S'82-M'84) was born in Cairo, Egypt, in 1954. He received the B.E. degree in electrical engineering from Cairo University in 1976, the M.A. degree in mathematics and the M.S. degree in computer science from the University of Pittsburgh in 1981, and the Ph.D. degree in computer science from the University of Pittsburgh in December 1983.

Since September 1989, he has been an Associate Professor of Computer Science at the University of Pittsburgh. Previously, he has been an Assistant Professor at Purdue University and at the University of Pittsburgh. His current research interests include fault-tolerant processor arrays, parallel algorithms, parallel architectures, and optical computing.

Dr. Melhem is a member of the Association for Computing Machinery, the IEEE Computer Society, and the International Society for Optical Engineering.



Ghil-Young Hwang (S'88-M'89) received the B.S. degree in physics from the Sogang University, Seoul, Korea, in 1981 and the M.S. degree in computer science from the University of Pittsburgh in 1985.

During 1986-1987, he conducted postgraduate studies at the University of Pittsburgh. His research interests include software engineering, parallel and distributed processing, and artificial intelligence. Currently, he is a Software Engineer at the Sukus Computer Company.

Mr. Hwang is a member of IEEE Computer Society.