

ÇOKLU ORTAM UYGULAMALARI İÇİN TEK/ÇİFT DUYARLIKLIL KAYAN NOKTA ÇARPICI TASARIMI

¹Metin Mete ÖZBİLEN

²Mustafa GÖK*

¹Bilgisayar Mühendisliği Bölümü, Mersin Üniversitesi,
Çiftlikköy, 33342, Mersin *e-posta*: mmozbilen@mersin.edu.tr
²Elektrik-Elektronik Mühendisliği Bölümü, Çukurova Üniversitesi,
Balcalı, 01330, Adana *e-posta*: musgok@cu.edu.tr

ABSTRACT

Modern graphic processors, multimedia processors, and general-purpose processors with multimedia extensions provide SIMD floating-point instructions. SIMD floating-point multiplication is commonly used in 2D and 3D applications, which mostly use single precision floating-point operands. Consequently, efficient single precision multiplier units are crucial for high performance systems. This paper introduces a novel method for performing two parallel single precision multiplications on a double precision floating point multiplier circuit. The proposed design uses approximately 10% more hardware and has 33% more delay compared to a conventional double precision floating-point multiplier. However, the proposed design is capable of executing two parallel single precision multiplications in one cycle.

Anahtar sözcükler: Çoklu ortam, Kayan-nokta sayılar, Çarpım

1. GİRİŞ

Çoklu ortam uygulamalarını çalıştırmaya yönelik donanımlar genel amaçlı sistemlerde standart hale gelmiştir [1]. Yeni üç boyutlu (3B) çoklu ortam uygulamaları yüksek duyarlılığın tersine performansı hedef almıştır [2,3]. Modern IEEE-754 [4] kayan-nokta tasarımları hem tek hem de çift-duyarlılıkta yüksek performansı hedeflemektedir [1 - 6].

IEEE-754 kayan nokta komutlarını destekleyen işlemciler hem tek hem de çift duyarlılıkta işlemleri gerçekleştirebilmektedir. Kayan nokta çarpımı en çok çoklu ortam işlemlerinde kullanılmaktadır. Özellikle

vektör çarpımlarını çalıştırmaya yönelik devreler, genel amaçlı işlemcilerin çoklu ortam uzantılarına (multimedia extensions) eklenmiştir. Bunlar arasında bazı popüler yonga üreticilerinin önerdikleri uzantılar şunlardır: Intel Pentium 3 mimarisi ile kullanmaya başladığı SSE ve SSE2 çoklu ortam uzantılarında paketlenmiş iki adet çift duyarlılıkta kayan nokta veya dört adet tek kayan nokta çarpımı yapılabilmektedir [5]. AMD'nin 3D Now çoklu ortam uzantısı da benzer bir şekilde aynı komutla iki tek duyarlılıkta kayan nokta çarpımını gerçekleştirmektedir [6]. PowerPC'nin Altivec uzantısı da dört adet tek duyarlılıkta çarpma işlemi yapan komutu içermektedir [7].

Sadece çift duyarlılıkta çarpımı destekleyen üniteler, genelde her iki duyarlılığı da destekleyen ünitelerden hem alan hem de kritik zaman açısından üstündür. Ancak bu donanımlarda tek duyarlılıkta çarpma yapmaya kalkışılınca performans oldukça düşmektedir. Bunun sebebi tek duyarlılıkta sayının önce çift duyarlılığa çevrimi sonra da yeniden tek duyarlılığa aktarılması ihtiyacıdır. Üstelik bir çift duyarlılıkta çarpıcı da her çevrimde yalnızca bir çarpım başlatılabilmektedir. Bu sorunları çözmek üzere [2]'de yapılan çalışmada her iki kipte çalışan IEEE-754 çarpıcısı sunulmuştur. Önerilen tasarımda bir çift duyarlılıkta kayan nokta çarpıcısı kullanılarak iki tek duyarlılıkta çarpımın gerçekleştirilmesi hedeflenmiştir. Ancak [2]'de önerilen çarpıcı da yeni bir çift duyarlılıkta çarpımın başlatılması için bir duraklama çevrimi gerekmektedir.

Bu makalede yeni bir yöntem izlenerek bir çift duyarlılıkta çarpıcı devreye iki tek duyarlılıkta çarpma yapma işlevselliği kazandırılmıştır. Bu sayede çok fazla verinin çarpıldığı uygulamalarda işlem

* Bu çalışma Çukurova Üniversitesi tarafından desteklenen *Kayan Noktalı Aritmetik Kullanan Çoklu Ortam Çarpma Devrelerinin Tasarlanması* başlıklı ve MMF2004BAP18 numaralı proje kapsamında gerçekleştirilmiştir.

performansı yaklaşık iki katına çıkabilecektir. Önerilen tasarım geleneksel bir çarpıcı ile karşılaştırılmıştır.

2. TEK/ÇİFT KAYAN NOKTA ÇARPICI

Bu kısımda bir çift duyarlıklı çarpıcı devresinde iki tek duyarlıklı çarpma işlemi eş zamanlı olarak gerçekleştiren tasarımımızı genel hatlarıyla sunacağız. Öncelikle hedeflediğimiz işlemi açıklayalım:

Şekil-1.(a)'da *IEEE-754* formatındaki çift duyarlıklı kayan nokta sayılar olan X ve Y operatörlerinin ve çarpımın sonucu olan Z sayısının üç 64-bit yazaç içerisindeki dizilişi görülmektedir. Şekil-1.(b)'de ise dört adet tek duyarlıklı *IEEE-754* biçimindeki A , B , C , ve D sayılarının sırasıyla iki 64-bitlik yazaçta dizilişleri görülmektedir. Üçüncü yazaçta ise A ve B sayılarının çarpımının sonucu üretilen H ve C ve D sayılarının çarpımının sonucu üretilen J sayılarının dizilişi görülmektedir.

	63	62	...	52	51	...	0
X	S_x	E_x		M_x			
Y	S_y	E_y		M_y			
Z	S_z	E_z		M_z			

(a) Çift duyarlıklı iki kayan nokta sayının yazaçlara yerleşimi.

	63	62...55	54...32		31	30...23	22...0
A	S_a	E_a	M_a	C	S_c	E_c	M_c
B	S_b	E_b	M_b	D	S_d	E_d	M_d
H	S_h	E_h	M_h	J	S_j	E_j	M_j

(b) Tek duyarlıklı dört kayan nokta sayının yazaçlara yerleşimi.

Şekil-1 Yazaç dizilişleri.

Şekil-1a'daki iki çift duyarlıklı X ve Y kayan nokta sayının çarpımı

$$E_z = E_x + E_y \quad (1)$$

$$M_z = M_x \times M_y \quad (2)$$

$$S_z = S_x \otimes S_y \quad (3)$$

denklemleri ile gerçekleştirilir. Şekil-1b'deki tek duyarlıklı A ve B sayılarının çarpım işlemi ve C ve D sayılarının çarpım işlemi

$$E_h = E_a + E_b, E_j = E_c + E_d \quad (4)$$

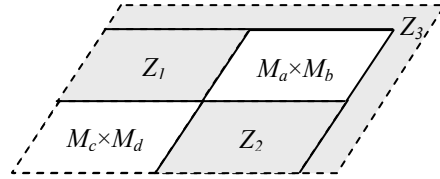
$$M_h = M_a \times M_b, M_j = M_c \times M_d \quad (5)$$

$$S_h = S_a \otimes S_b, S_j = S_c \otimes S_d \quad (6)$$

denklemleri ile gerçekleştirilir. Bizim tasarımımızda bu iki tek duyarlıklı kayan nokta çarpımın eş zamanlı gerçekleştirilmesi hedeflenmiştir.

İki işaretli sayının çarpımını gerçekleştiren teknik [9] da bahsedilmiştir. Bu teknik kullanılarak iki tek duyarlıklı kayan noktalı çarpım için üretilen matrisler,

bir çift duyarlıklı çarpım için üretilen matrisin içine yerleştirilebilir. Şekil-2 kayan noktalı sayılar için bu tekniğin nasıl uygulandığını gösteren bir diyagram gösterilmiştir. Bu diyagramda kesik çizgiler arasında gösterilen büyük matris içindeki tüm çarpım bitleri çift duyarlıklı çarpım kipinde üretilmektedir. Tek duyarlıklı çarpım kipinde ise, bu matrisin içindeki gölgeli alanlardaki çarpım bitleri (Z_1, Z_2, Z_3) sıfıra eşitlenmektedir. Sağ üst gölgesiz alandaki bitler M_a ve M_b mantislerinin çarpımı sonucu üretilmiştir. Sol alt gölgesiz alandaki bitler M_c ve M_d mantislerinin çarpımı sonucu üretilmiştir



Şekil-2 Tek duyarlıklı paralel çarpımların kısmi çarpım matrisine yerleştirilmesi

Şekil-2'deki gösterilen Z_1 ve Z_2 ile etiketlenmiş bölgelerdeki bitleri üretmek için

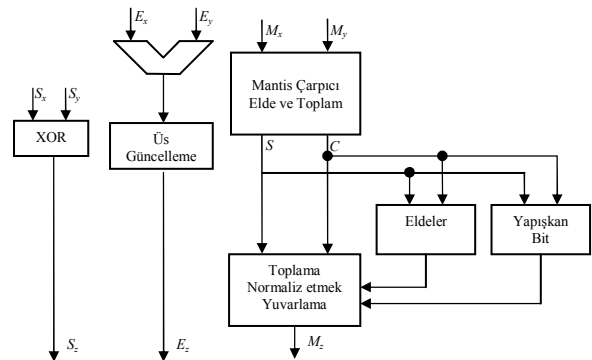
$$\hat{b}_j = s \cdot b_j \text{ ve } p_{ij} = a_i \cdot \hat{b}_j \quad (7)$$

mantık denklemleri kullanılmaktadır. Diğer bölgelerdeki bitler ise basit VE işlemi ile elde edilir.

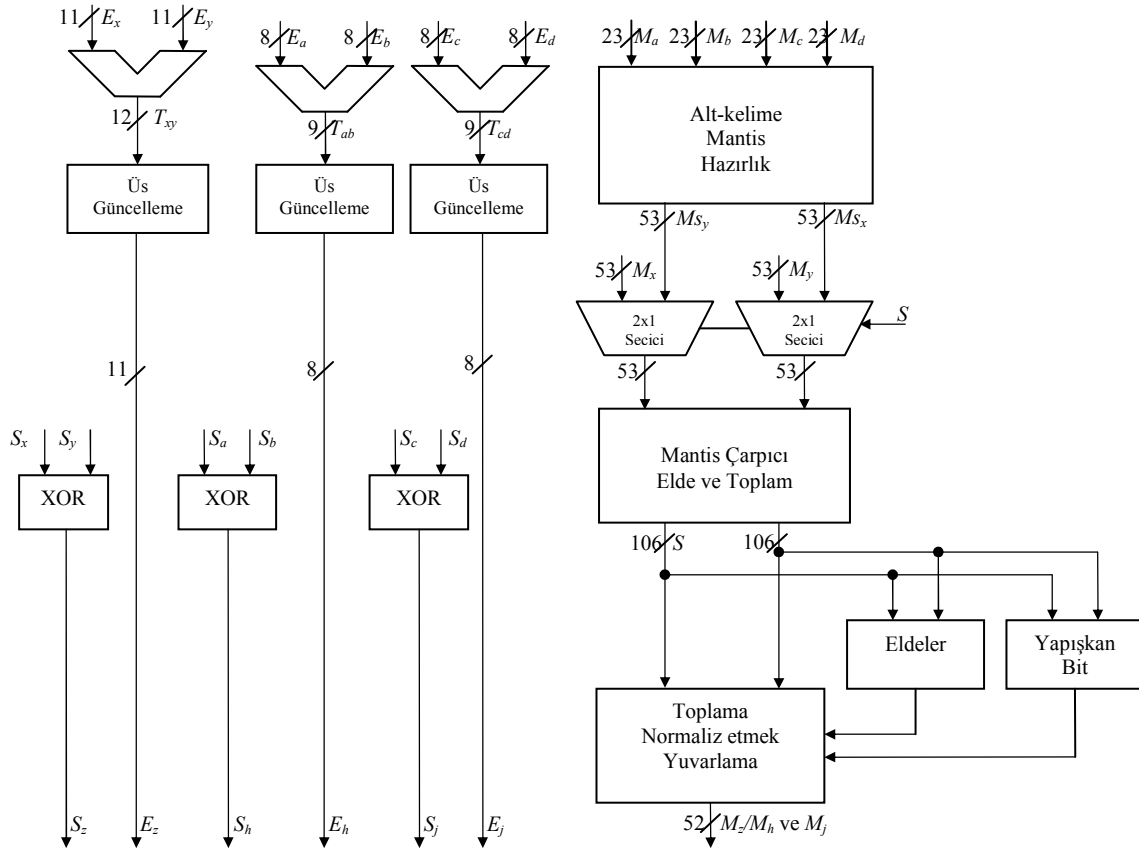
$$p_{ij} = a_i \cdot b_j \quad (8)$$

Böylece $s=0$ iken matriste tek duyarlıklı iki çarpmaya uygun olarak bitleri üretilirken, $s=1$ iken matriste çift duyarlıklı çarpmaya uygun bitler üretilmiştir.

Hızlı çarpıcılarda üretilen çarpım bit matrisi herhangi bir ağaç indirgeme yöntemi ile iki vektöre indirgenir; daha sonra bu iki vektör bir elde yayımlı toplayıcı (carry-propagate adder) yardımı ile toplanarak sonuç üretilir. Devremizin iki kipinde de aynı toplayıcı devre kullanılabilir.



Şekil-3 Klasik kayan nokta çarpıcının blok diyagramı.



Şekil-4. Önerilen tek/çift duyarlıklı kayan nokta çarpıcının blok diyagramı.

Tasarımımızın getirdiği eklentileri belirtmek için önce, klasik çarpma algoritmasını uygulayan bir çift duyarlıklı kayan nokta çarpıcı devresi Şekil-3'de gösterilmiştir.

Şekil-4'de de bu çarpıcı üzerinde yapılan basit eklentilerle elde edilen tek/çift duyarlıklı çarpıcı tasarımının gerçekleştirimi gösterilmektedir. Önerdiğimiz eklentiler genel olarak her tür çift duyarlıklı kayan nokta çarpıcı devreye konularak, uygulanan çarpıcıya iki tek duyarlıklı çarpım işlevselliği katılabilir.

Şekil 4'de önerilen çarpıcının veri akışını ve blokların işlevlerini kısaca şu şekilde açıklayabiliriz:

- S , kontrol sinyali '0' ise çift duyarlıklı çarpma gerçekleştirilmekte, aksi halde iki tek duyarlıklı çarpma gerçekleştirilmektedir.
- Bir adet 11-bitlik toplayıcı çift duyarlık kipinde iki kayan nokta sayının üslerini ve iki adet 8-bitlik toplayıcı ise tek duyarlık kipinde dört kayan nokta sayının üslerini paralel olarak toplamaktadır.
- Üs güncelleme üniteleri üslerin toplamı sonucu fazladan eklenen sapma (bias) miktarlarını çıkarmaktadır.
- Mantıs Hazırlama bloğunda, Denklem (7)'de gösterilen mantık işlemleri yapılmaktadır.

- Mantıs çarpıcı mantısları çarpmakta ve normalize edilmemiş ve yuvarlanmamış sonucu üretmektedir.
- Çarpımın veya çarpımların sonuçlarının işareti XOR kapıları ile belirlenmektedir.
- En sondaki blok çarpımları yuvarlamak ve normalize etmektedir. Tasarımda yalnızca IEEE-754 sıfıra doğru yuvarlama kipini kullandık; pratikte bu düşük çözünürlükteki bitlerin kırılması demek olduğu için fazladan bir yuvarlama devresi kullanılmamıştır.

3. SENTEZ SONUÇLARI

Bu bölümde önerdiğimiz tek/çift duyarlıklı kayan nokta çarpıcı ve klasik bir çift duyarlıklı kayan nokta çarpıcının sentez yoluyla karşılaştırılması sonucu elde edilen verileri sunuyoruz. Karşılaştırılan iki devre yapısal VHDL kullanılarak modellenmiştir. Yukarıda açıklandığı üzere iki modelde de yuvarlama kırılarak yapılmıştır.

Devre sentezi bir uygulamaya yönelik bütünlük devre teknolojisi olan TSMC 0.18 mikron kütüphanesi ve LeonardaSpectrum LS2005b yazılımı kullanılarak yapılmıştır. Her iki devre de kritik zamanın azaltılması hedefi ile yüksek optimizasyon eforu seçeneği kullanılarak sentezlenmiştir. Tablo-1 de gösterilen

değerler alan için kapı sayıları, zaman için nanosaniye (ns) cinsindedir.

Tasarım	Alanı	Kritik süresi
Klasik çift duyarlılık	25175	4.10 ns
Tek/çift duyarlılık	27566	5.49 ns

Tablo-1 Klasik çift duyarlılık ve önerilen tek/çift duyarlılık kayan nokta çarpıcıların karşılaştırılması.

Bu değerlere göre tek/çift duyarlılık çarpma yapabilen tasarımı yaklaşık %9.49 daha fazla alan kaplamakta %34 oranında fazla kritik süreye sahiptir. Modern işlemcilerde kullanılan kayan nokta çarpıcıları klasik kayan nokta çarpım algoritmasını kullanmamaktadır. Daha büyük bir yapıda ve boru hatlı (pipelined) tasarımlara sahip olan bu tür çarpıcılara önerdiğimiz değişiklikler yapıldığı takdirde yüzde alan artışı %5 in altına düşecektir, ayrıca ekstra süre artışı da boru hattının kademelerinde eritilebilecektir. Süre farkını azaltacak diğer bir faktörde yuvarlama ünitelerinin kullanılması olacaktır.

4. SONUÇ

Bu çalışmada bir *IEEE-754* tek/çift duyarlılık kayan nokta çarpıcı tasarımı sunulmuştur. Tasarımımızda kullanılan yöntemler daha uzun kayan nokta veri yollarına da rahatlıkla uygulanabilir.

Bu çalışmada sunulan tasarım yönteminin literatürdeki benzerlerinden önemli bir üstünlüğü her tür çarpıcı tasarımına kolaylıkla uyarlanabilmesidir. Tasarımımızı sentez yoluyla klasik kayan nokta çarpıcı ile karşılaştırdık. Karşılaştırma sonuçları tasarımımızın klasik çarpıcıya göre yaklaşık %10 daha fazla alan kapladığını, ve kritik zaman artışının ise bir kaç kapı gecikmesi ile sınırlı olduğu göstermektedir. Ancak tasarımı iki çalışma kipinde çalışabilmekte ve özellikle tek duyarlık kipinde aynı anda iki çarpım yapabilmektedir. İki ve üç boyutlu uygulamaların gittikçe artan kullanımı sunduğumuz türde çarpıcı tasarımların önem kazanacağını göstermektedir.

Bu çalışma tek/çift duyarlılık vektör çarpımlarına yönelik bir çarpıcı-toplayıcı devresi tasarımıyla genişletilecektir. Ayrıca ilerideki çalışmalarımızda bütün *IEEE-754* yuvarlama kiplerini destekleyecek eklentiler de yapılacaktır.

KAYNAKLAR

- [1] Lappalainen, V., Hämäläinen, T. D., Liuha, P., Overview of Research Efforts on Media ISA Extensions and Their Usage in Video Coding, *IEEE Transactions On Circuits And Systems For Video Technology*, Vol 12, 2002, No. 8.
- [2] Even, G., Mueller, S.M., Seidel, P.-M., A dual mode IEEE multiplier, *Proceedings., Second*

Annual IEEE International Conference, pp. 282-289, 1997.

- [3] Lee, R.B., Multimedia extensions for general-purpose processors, *Signal Processing Systems, 1997. SIPS 97 - Design and Implementation., 1997 IEEE Workshop*, pp 9-23.
- [4] ANSI-IEEE Standard 754-1985: IEEE Standard for binary floating-point arithmetic, 1985.
- [5] IA-32 Intel® Architecture Software Developer's Manual, <http://download.intel.com/design/Pentium4/manuals/25366520.pdf>, 2006
- [6] Advanced Micro Devices, "AMD64 Architecture Programmers Manual", White Paper, http://www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/26569.pdf, 2006
- [7] Diefendorff, K.; Dubey, P.K.; Hochsprung, R.; Scale, H., AltiVec extension to PowerPC accelerates media processing, *IEEE Micro*, Volume 20, 2000, Issue 2, pp. 85-95.
- [8] Jennings, M. D., Conte, T. M., Subword Extensions for Video Processing on Mobile Systems, *IEEE Concurrency*, Vol.6, 1998, No. 3, pp 13-16.
- [9] Gök M., Krithivasan, S., Schulte, M.J., Designs for Subword-Parallel Multiplications and Dot Product Operations, *Workshop on Application Specific Processors*, pp. 27-31, 2004