

# DALGACIK TEPELERİ VE DESTEK VEKTÖR MAKİNELERİ İLE MÜZİK ÇALGISI SINIFLANDIRMA

M. Erdal ÖZBEK<sup>1</sup>, Nalan ÖZKURT<sup>2</sup> ve F. Acar SAVACI<sup>1</sup>

<sup>1</sup> Elektrik-Elektronik Mühendisliği Bölümü, İzmir Yüksek Teknoloji Enstitüsü  
Urla, 35430, İzmir. e-posta: {erdalozbek,acarsavaci}@iyte.edu.tr

<sup>2</sup> Elektrik-Elektronik Mühendisliği Bölümü, Dokuz Eylül Üniversitesi  
Buca, 35160, İzmir. e-posta: nalan.ozkurt@eee.deu.edu.tr

## ABSTRACT

*The automatic classification of musical instruments is studied using the feature vectors obtained from wavelet ridges. The wavelet ridges identifying the energy localizations of isolated note samples are computed with a simplified method using singular value decomposition. A feature vector which is composed of wavelet ridges, their corresponding frequencies and number of ridges for each analysis window is calculated from each note sample. All possible two-class classifications from five woodwind instruments are performed using support vector machines with linear, polynomial and radial basis function kernels. The correct classification rates for each instrument are obtained. Our best result is achieved with 82% for Bassoon and Oboe pair.*

**Anahtar Sözcükler:** Müzik Çalgısı Sınıflandırma, Dalgacık Tepeleri, Destek Vektör Makineleri.

## 1. GİRİŞ

Son yıllarda müzik işaretlerinin ayrıştırılması popüler bir araştırma konusu olmuştur [1]. Müzik işaretlerinin ayrıştırılmasının arkasında çoğunlukla aynı anda çalınan çalgıların ve çaldıkları notaların tanınarak ayrıştırılması gelmektedir. Çalgı ve notaların çok sayıda olası bileşimleri nedeniyle bu işlem oldukça karmaşıktır ve birçok çalgı tarafından çalınan bir müzikten nota sembollerinin elde edilmesi henüz mümkün olmamıştır.

Müzik çalgısı tanıma ya da sınıflandırma amacıyla yapılan çalışmalar ilk olarak ayrıştırılmış notaları kullanmıştır. Çalgı seslerinden elde edilen bu nota örnekleri kullanılarak sınıflandırma işlemi temel olarak, nota örneklerinden özniteliklerin çıkarılması ve değişik şekillerde oluşturulan öznitelik vektörünü kullanarak

bir öğrenme algoritması yardımıyla nota ya da çalgı sınıflarına ayrıştırılmasından oluşmaktadır [2].

Ses tanıma sistemlerinde kullanılan teknikler yardımıyla problemin çözülmesi için birçok araştırma yapılmış, bunların ve genel işaret işleme yöntemlerinin kullanılmasıyla çok geniş bir öznitelik kümesi elde edilmiştir. Bunlara örnek olarak özilinti katsayıları, sıfır geçiş sayısı, izgesel akı, izgesel kitle merkezi ve mel frekans kepsral katsayılarını verebiliriz [3]. Bu çalışmada ise dalgacık tepeleri kullanılarak öznitelik vektörleri elde edildikten sonra destek vektör makineleri kullanılarak sınıflandırma yapılmıştır.

İzleyen bölümde dalgacık tepelerinin elde edilme yöntemi anlatılacaktır. Üçüncü bölüm destek vektör makineleri hakkında özet bilgi içermekte ve dördüncü bölüm nota işaretlerinden öznitelik vektörünün oluşturulması için yapılan benzetim çalışmaları ve elde edilen sınıflandırma başarımlarını sonuçlarını vermektedir. Son bölümde ise sonuçlar özetlenip, gelecekteki çalışmalar tartışılacaktır.

## 2. DALGACIK TEPELERİ

Sürekli dalgacık dönüşümünde işaret  $s(t)$ ,  $\Psi$  ile gösterilen ve temel dalgacık adı verilen işlevin sıkıştırılması ve kaydırılması ile oluşturulan dalgacık uzayına yansıtılır ve dalgacık katsayıları aşağıdaki şekilde elde edilir:

$$W_s(a, b; \Psi) \triangleq \int_{-\infty}^{\infty} s(t) \Psi_{a,b}^*(t) dt. \quad (1)$$

Burada  $a$  ve  $b$  sırasıyla sıkıştırma ve kaydırma katsayılarıdır. Sıkıştırılmış ve kaydırılmış dalgacıklar ise

$$\Psi_{a,b}(t) \triangleq \frac{1}{a} \Psi \left( \frac{t-b}{a} \right), \quad a \in \mathbb{R}^+, \quad b \in \mathbb{R}, \quad (2)$$

şeklinde elde edilir. Bu çalışmada temel dalgacık olarak aşağıda ifadesi verilen Morlet temel dalgacı

<sup>†</sup> Bu çalışma TÜBİTAK-EEEAG tarafından desteklenmektedir.

<sup>‡</sup> Nota örneklerinin hazırlanması ilk yazarın Paris'teki Supelec-LSS laboratuvarında bulunduğu sürede yapılmıştır.

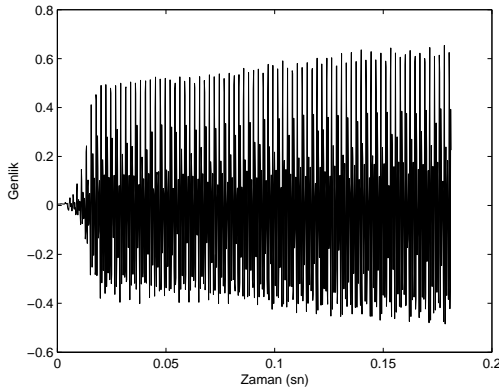
kullanılmıştır.

$$\psi(t) = \frac{1}{\sqrt{2\pi}} e^{j\omega_0 t} e^{-t^2/2} \quad (3)$$

Burada  $\omega_0$  dalgacığın merkez frekansıdır.

Bir işaretin anlık frekansı basitçe işaretin fazının türevi olarak tanımlanabilir ve böylece işaretler temel frekansı değişen frekansı modüle edilmiş işaretler kullanılarak modellenabilir. Dalgacık tepeleri, işaretin fazının türevinin sıfır olduğu noktaları bulmak amacıyla durağan faz kuramını kullanarak [4] ya da skalogramın yerel tepe noktalarının bulunması olarak ifade edilen basit yöntemle [5] bulunabilir. Ancak bu yöntemler gürültüye karşı çok duyarlıdır. Gürültülü işaretlerin dalgacık tepelerinin elde edilmesi için istatistiksel yöntemlere dayanan bazı algoritmalar önerilmiştir [6, 7]. Bu yöntemlerde işlem yükünün fazla olması nedeniyle alternatif olarak [8]'de önerilen tekil değer ayrıştırma yöntemi kullanılmıştır. Skalogramda enerjinin yerel olarak yoğunlaştığı bölgeler gürültüden etkilenen enerji bileşenleri ile tekil değer ayrıştırma yöntemi ile ayrılır. Skalogram tekil değerlerine ayrıştırıldığı zaman tekil değerler azalan sırada olacağı için enerjinin çoğu ilk sıradaki tekil değerlerdedir. Toplamsal beyaz gürültü durumunda küçük tekil değerler büyüklere oranla gürültüden daha çok etkilendiği için, küçük tekil değerlerin ayrılıp yaklaşık skalogram matrisinin sadece büyük tekil değerler ile geri elde edilmesi ile gürültünün etkisi azaltıldığı gibi işlem yükü de azaltılmıştır.

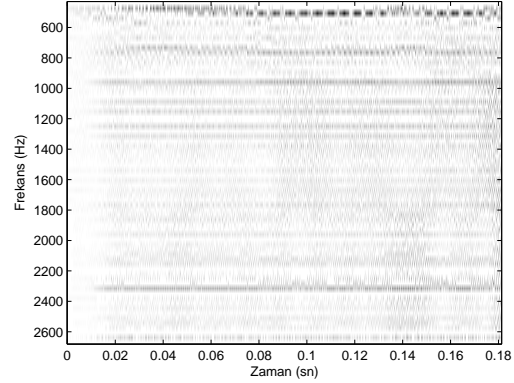
Şekil 1'de, Obua çalgısının 440Hz frekansındaki 'A4' olarak da adlandırılan 'La' notası örneğinden bir pencere gösterilmektedir.



Şekil 1: Obua La notası örneği

Çalışmamızda Şekil 1'deki gibi nota örneklerinden elde edilen pencereler boyutunda hesapladığımız yeni yaklaşık skalogramı kullanarak dalgacık tepeleri elde edilmektedir. Verilen örneğin dalgacık tepelerini gösteren yeni skalogramı Şekil 2'de gösterilmektedir.

Enerjinin yoğun olduğu frekans değerleri beklendiği gibi 440Hz civarında görünmektedir.



Şekil 2: Obua La notası dalgacık tepeleri

### 3. DESTEK VEKTÖR MAKİNELERİ

Destek vektör makinelерinin temelleri istatistiksel öğrenme kuramına dayanılarak Vapnik tarafından atılmıştır [9]. Bu kurama göre belirli bir öğrenme işlemi için sınırlı sayıda veri verildiğinde en iyi genelleştirme, sınıflandırma işlevinin kapasitesinin eğitim kümesinin boyutu ile uyumlu olduğu durumda elde edilir. İlk uygulama en büyük aralık sınıflandırıcısı adıyla eğitim örnekleri ile sınıf ayırıcı sınırın arasındaki uzaklığı en büyükleen bir eğitim algoritması olarak verilmiştir [10].  $n$  boyutlu eğitim örnekleri  $\mathbf{x}$  ve sınıf etiketi  $y$  ( $y_i \in \{-1, 1\}, i = 1, 2, \dots, n$ ) ile verildiğinde algoritma  $\mathbf{w} \cdot \mathbf{x} + b = 0$  ile verilen en iyi ayırıcı yüzeyi bulmaya çalışır. Bu durumda

$$y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0, \quad \forall i \quad (4)$$

şeklinde yazılan optimizasyon problemi en yakın eğitim örneği ile ayırıcı yüzey arasındaki uzaklık  $2/\|\mathbf{w}\|$  en büyük olacak şekilde kısıtlandırılır. Ayırıcı yüzeye en yakın eğitim örnekleri, tüm eğitim örneklerinin küçük bir alt kümesini oluşturur ve destek vektörleri olarak adlandırılır [11].

En büyük aralık sınıflandırıcısı basittir ve doğrusal ayrıştırılabilen sınıflandırma için önerilmiştir. Ancak veriler doğrusal olarak sınıflandırılmadığında giriş uzayı  $X$ 'ten verilerin doğrusal olarak sınıflandırılacağı daha büyük (genellikle sonsuz) boyutlu  $\mathcal{H}$  uzayına  $\phi$  dönüşümü ile geçişi sağlayan bir çekirdek işlevi,  $K$ , kullanılmaktadır.

$$K(\mathbf{x}, \mathbf{z}) = \phi(\mathbf{x}) \cdot \phi(\mathbf{z}), \quad \forall \mathbf{x}, \mathbf{z} \in X \quad (5)$$

$\mathcal{H}$ 'daki iç çarpım, çekirdek hilesi olarak bilinen bir değişim sayesinde  $\phi$  işlevinin açık ifadesine gerek kalmadan hesaplanmaktadır. Böylece bir uzayda

iç çarpım olup olmadığını belirleyen Mercer şartını sağladığı sürece en iyi ayırıcı yüzeyi oluşturmak için herhangi bir işlev kullanılabilir [12]. En sık kullanılan çekirdek işlevleri,

$$\begin{aligned} K(\mathbf{x}, \mathbf{z}) &= (\mathbf{x} \cdot \mathbf{z}) && : \text{doğrusal} \\ K(\mathbf{x}, \mathbf{z}) &= (\mathbf{x} \cdot \mathbf{z} + 1)^d && : \text{polinom} \\ K(\mathbf{x}, \mathbf{z}) &= \exp\left(-\frac{\|\mathbf{x}-\mathbf{z}\|^2}{2\sigma^2}\right) && : \text{radyal taban işlevi} \end{aligned}$$

olarak verilebilir. Ayrıca belirli amaçlar ve uygulamalar için bir çok çekirdek işlevi bulunmaktadır [13].

Destek vektör makineleri esas olarak iki sınıf sınıflandırması için tasarlanmış olmasına rağmen bire-bir ve bire-tüm metotları kullanılarak çok sınıflı sınıflandırma gerçekleştirilebilmektedir. Her ikisi de çok sınıflandırma problemini iki sınıf sınıflandırma problemlerinin toplamı olarak ele almaktadır.  $k$  sınıf sınıflandırma için, bire-tüm metodu bir sınıf ile kalan  $k-1$  sınıf arasında yüzeyi oluştururken, bire-bir metodunda her olası sınıf çifti arasındaki  $\frac{k(k-1)}{2}$  sayısındaki yüzeyler oluşturulmaktadır. Her iki yöntemde de belirlenen bir seçme kuralı ile karar verilmektedir.

#### 4. BENZETİM ÇALIŞMALARI VE SINIFLANDIRMA BAŞARIMI

Bu çalışmada Iowa Üniversitesi nota örnekleri [14] kullanılmıştır. Bu örnekler tek kanallı, 16 bit çözünürlüklü ve 44100Hz örnekleme frekansına sahiptir. Grup halinde bulunan notalar tek tek ayrıştırılarak<sup>‡</sup> herbir nota için öznitelikler hesaplanmıştır. Nota örnekleri başlangıç ve sonlarındaki sessiz bölümler kullanılan bir eşik değeri yardımıyla çıkartılarak sonraki hesaplamalara hazır hale getirilmiştir.

Hesaplama kullanılan sürekli dalgacık dönüşümünün olası tüm frekans değerleri için hesaplanması uzun sürmektedir. Bu nedenle işaretin tümü yerine enerjinin yoğunlaştığı frekans değerleri çevresinde sürekli dalgacık dönüşümü uygulanmıştır. Bu amaçla ilk olarak, işaretin hızlı Fourier dönüşümü alınarak işaretin frekans bileşenleri bulunmuştur. Dalgacık dönüşümü sadece bu frekanslar için hesaplanmış, bu sayede benzetimlerde süre indirimi sağlanmıştır.

Dalgacık tepeleri, hesaplamalarda kolaylık sağlanması açısından örnekleme frekansı azaltıldıktan sonra, %50 oranında örtüşen pencereler kullanılarak elde edilmiştir. Her penceredeki en büyük 5 dalgacık tepesi enerji değeri, bu değerlerin elde edildiği frekans değeri ve her penceredeki dalgacık tepeleri sayısından oluşan bir vektör oluşturulmuştur. Nota işareti süresince elde edilen bu vektörlerin ortalama değeri öznitelik vektörü olarak seçilmiştir.

Benzetim sonuçları beş üflemlerli çalgıdan (Alto Saksofon, Bas Klarinet, Fagot, Flüt, Obua) olası tüm ikili grupları iki sınıf sınıflandırıcı destek vektör makineleri ile sınıflandırarak elde edilmiştir. Çekirdek işlevi olarak doğrusal çekirdeğin yanı sıra, polinom

çekirdeğinin farklı dereceleri ( $d$ ) ve radyal taban işlevinin değişen  $\sigma$  değerleri için sınıflandırma gerçekleştirilmiştir.

Destek vektör makineleri için nesne tabanlı Spider [15] yazılımı kullanılmış, her bir çalgı için elde edilmiş öznitelik vektörlerinin yarısı eğitime aşamasında kalan yarısı da test aşamasında kullanılmıştır. Tablo 1’de doğrusal çekirdek kullanılarak elde edilen başarımlar oranları verilmiştir.

Tablo 1: Doğrusal çekirdek sınıflandırma başarımları

| Başarım oranı (%) | Saksofon |              | Fagot | Flüt | Obua |
|-------------------|----------|--------------|-------|------|------|
|                   | Alto     | Bas Klarinet |       |      |      |
| Alto Saksofon     | -        | 54.3         | 55.5  | 73.1 | 74.2 |
| Bas Klarinet      | 54.3     | -            | 54.3  | 59.9 | 59.5 |
| Fagot             | 55.5     | 54.3         | -     | 57.2 | 75.0 |
| Flüt              | 73.1     | 59.9         | 57.2  | -    | 60.6 |
| Obua              | 74.2     | 59.5         | 75.0  | 60.6 | -    |

Sonuçlardan görüldüğü üzere Fagot ve Obua ikilisinin birlikte olduğu durumda iki çalgının doğru olarak sınıflandırılma oranı %75 olmasına rağmen diğer çalgılar için bu oran yakalanamamıştır. Sonuçları Tablo 2’de verilen ikinci dereceden ( $d=2$ ) polinom çekirdek kullanıldığında ise %82’nin üzerinde bir oran elde edilmiş ve tüm diğer çalgıların sınıflandırılma oranları da daha yüksek bulunmuştur.

Tablo 2: 2. dereceden polinom çekirdek sınıflandırma başarımları

| Başarım oranı (%) | Saksofon |              | Fagot | Flüt | Obua |
|-------------------|----------|--------------|-------|------|------|
|                   | Alto     | Bas Klarinet |       |      |      |
| Alto Saksofon     | -        | 67.1         | 78.4  | 66.4 | 70.1 |
| Bas Klarinet      | 67.1     | -            | 65.9  | 72.5 | 63.6 |
| Fagot             | 78.4     | 65.9         | -     | 68.2 | 82.1 |
| Flüt              | 66.4     | 72.5         | 68.2  | -    | 67.9 |
| Obua              | 70.1     | 63.6         | 82.1  | 67.9 | -    |

Benzetimler artan polinom çekirdek dereceleri için de gerçekleştirilmiş, ancak polinom derecesi artırıldığında örnek olarak dördüncü dereceden polinom için Tablo 3’de verilen sonuçlar ışığında başarımın artmadığı gözlenmiştir.

Tablo 3: 4. dereceden polinom çekirdek sınıflandırma başarımları

| Başarım oranı (%) | Saksofon |              | Fagot | Flüt | Obua |
|-------------------|----------|--------------|-------|------|------|
|                   | Alto     | Bas Klarinet |       |      |      |
| Alto Saksofon     | -        | 64.0         | 65.8  | 68.3 | 61.9 |
| Bas Klarinet      | 64.0     | -            | 62.0  | 57.7 | 60.3 |
| Fagot             | 65.8     | 62.0         | -     | 67.6 | 60.7 |
| Flüt              | 68.3     | 57.7         | 67.6  | -    | 63.6 |
| Obua              | 61.9     | 60.3         | 60.7  | 63.6 | -    |

Benzetimler bir diğer çekirdek olan radyal taban işlevi kullanılarak da gerçekleştirilmiş, farklı  $\sigma$  değer-

leri için sınıflandırma başarımı incelenmiştir. Değişen  $\sigma$  değerlerinde farklı başarımlar elde edilse de, örnek olarak  $\sigma = 1$  için elde edilen sonuçlar Tablo 4’de verildiği gibi başarımın ikinci dereceden polinom kullanıldığında elde edilen sonuçlardan daha iyi olmadığını göstermektedir.

Tablo 4: Radyal taban işlevi çekirdek sınıflandırma başarımı

| Başarımlar oranı (%) | Alto<br>Saksofon | Bas<br>Klarnet | Fagot | Flüt | Obua |
|----------------------|------------------|----------------|-------|------|------|
| Alto Saksofon        | -                | 62.2           | 67.1  | 64.4 | 59.9 |
| Bas Klarnet          | 62.2             | -              | 61.2  | 62.1 | 58.7 |
| Fagot                | 67.1             | 61.2           | -     | 59.4 | 63.4 |
| Flüt                 | 64.4             | 62.1           | 59.4  | -    | 67.3 |
| Obua                 | 59.9             | 58.7           | 63.4  | 67.3 | -    |

## 5. SONUÇLAR VE DEĞERLENDİRME

Bu bildiri müzik çalgılarının dalgacık tepeleri kullanılarak sınıflandırılmasına çalışılmıştır. Bu amaçla veritabanından ayrıştırılmış nota örneklerine sürekli dalgacık dönüşümü uygulanmış ancak sürekli dalgacık dönüşümünün hesaplama yoğunluğunu azaltmak amacıyla da işaretin hızlı Fourier dönüşümünden elde edilen frekans değerleri çevresine yoğunlaştırılmıştır. Elde edilen enerji yoğunluğu skalogramına daha sonra uygulanan tekil değer ayrıştırması ile hem gürültü bileşenleri ayıklanmış hem de işlem yükü azaltılarak enerjinin yoğunlaştığı dalgacık tepeleri elde edilmiştir. Dalgacık tepelerinin enerji değerleri ve bu değerleri aldıkları frekans ile pencere başına düşen dalgacık tepesi sayısından oluşan öznitelik vektörü oluşturulmuş ve bu vektörlerin uygulanan pencereler toplamından elde edilen ortalama değeri, ayrıştırılmış her bir nota için sınıflandırma amacıyla kullanılmıştır.

Beş üflemeli çalgı için iki sınıflı destek vektör makineleri kullanılarak sınıflandırma yapılmıştır. Doğrusal, polinom ve radyal taban çekirdek işlevleri kullanılarak sonuçlar elde edilmiştir. Sonuçlar sınıflandırmada kullanılan çekirdek işlevinin önemini ortaya koymuştur. Fagot ve Obua için %82’nin üzerinde bir oran elde edilmiş ve bu oran, dalgacık tepelerinden elde edilen öznitelik vektörünün sınıflandırma için kullanılabilirliğini göstermiştir. Ancak bu oranın diğer çalgılarda ve farklı çekirdek işlevleri kullanıldığında elde edilememiş olması, öznitelik vektörlerinin yanı sıra kullanılan çekirdek işlevlerinin seçiminin de önemini ortaya koymuştur.

Çalışmada sonraki aşama olarak farklı öznitelik vektörlerinin elde edilerek karşılaştırma yapılması ve bu sayede en başarılı öznitelik vektörünün elde edilmeye çalışılması düşünülmektedir. Bunun yanında, iki sınıf sınıflandırıcı yerine de bire-bir veya bire-tüm metotları kullanılarak yapılacak çok sınıflı

sınıflandırma sonuçlarının elde edilmesi planlanmıştır.

## KAYNAKLAR

- [1] A. P. Klapuri, “Automatic Music Transcription as We Know it Today”, *Journal of New Music Research*, Vol. 33, No. 3, pp. 269-282, 2004.
- [2] P. Herrera-Boyer, G. Peeters, S. Dubnov, “Automatic Classification of Musical Instrument Sounds”, *Journal of New Music Research*, Vol. 32, No. 1, pp. 3-21, 2003.
- [3] A. Eronen, “Automatic musical instrument recognition”, Yüksek lisans tezi, Tampere University of Technology, April 2001.
- [4] C. Chandre, S. Wiggins, T. Uzer, “Time-frequency analysis of chaotic systems”, *Phys. D*, Vol. 181, pp. 171-196, 2003.
- [5] M. Todorovska, “Estimation of instantaneous frequency of signals using the continuous wavelet transform”, Univ. of Southern California, Los Angeles, Tech. Rep. CE 01-07, 2001.
- [6] R. Carmona, W. Hwang, “Characterization of signals by the ridges of their wavelet transforms”, *IEEE Trans. Signal Processing*, Vol. 45, No. 10, pp. 2586-2590, Oct. 1997.
- [7] R. Carmona, W. Hwang, B. Torresani, “Multiridge detection and time-frequency reconstruction”, *IEEE Trans. Signal Processing*, Vol. 47, No. 2, pp. 480-492, Feb. 1999.
- [8] N. Özkurt, F. A. Savacı, “Determination of Wavelet Ridges of Nonstationary Signals by Singular Value Decomposition”, *IEEE Trans. on Circuits and Systems-II: Express Briefs*, Vol. 52, No. 8, pp. 480-485, August 2005.
- [9] V. N. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, 1998.
- [10] B. E. Boser, I. M. Guyon, V. N. Vapnik, “A Training Algorithm for Optimal Margin Classifiers”, *5<sup>th</sup> Conference on Computational Learning Theory*, pp. 144-152, 1992.
- [11] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge University Press, 2000.
- [12] C. Cortes, V. Vapnik, “Support Vector Networks”, *Machine Learning*, Vol. 20, pp. 273-297, 1995.
- [13] J. Shawe-Taylor, N. Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, 2004.

[14] The University of Iowa Electronic Music Studios.  
<http://theremin.music.uiowa.edu>

[15] Spider: Object-orientated machine learning  
library.  
<http://www.kyb.tuebingen.mpg.de/bs/people/spider>