

Modelado de gran señal de transistores para alta frecuencia (HEMTs)

Wenceslao Rivas*

Resumen

En este trabajo se presenta una solución relativamente sencilla y precisa al problema de simular el comportamiento de un determinado HEMT cuando éste trabaja en una aplicación de gran señal. La solución consiste en el uso de un “modelo por tablas”, es decir, que cada parámetro está definido por una lista de valores que dependen de V_{gs} y V_{ds} . Estos valores se han de determinar por medio de un procedimiento que toma en consideración algunos problemas técnicos típicos en este tipo de aplicaciones. Como comprobación de la utilidad del modelo se han comparado los resultados obtenidos del análisis realizado por medio de un programa de simulación a un mezclador para 10 GHz., donde para describir al HEMT se ha usado el modelo presentado, con los resultados obtenidos de mediciones hechas al mezclador real.

1. Generalidades: conceptos básicos de modelado y no-linealidad

La necesidad de definir un modelo para un dispositivo electrónico surge cuando se desea diseñar un circuito, para cuyo análisis teórico se requiere que cada elemento físico que compone el circuito (resistencias, capacitancias, elementos semiconductores) cuente con una representación matemática (modelo) que permita simular la función de dicho elemento bajo las condiciones de operación del circuito. En la actualidad el análisis de circuitos se lleva a cabo por medio de programas de simulación, los cuales cuentan en las llamadas librerías con los modelos de los dispositivos electrónicos más comunes.

Toca ahora diferenciar lo que es el llamado “modelo de pequeña señal” y el “modelo de gran señal” de un dispositivo (p.ej., de un transistor). Se tiene primero el hecho de que por lo general los dispositivos semiconductores tienen un comportamiento no lineal de corriente-voltaje; sin embargo, para el caso en que las variaciones de señal, a que se encuentra sometido son pequeñas, este comportamiento no lineal suele aproximarse a un comportamiento lineal; luego si el dispositivo se comporta como un sistema lineal significa que puede encontrarse un circuito equivalente compuesto de elementos pasivos que simulen el comportamiento de aquel, en donde a este circuito equivalente se le conoce como el modelo de pequeña señal del dispositivo. Por otra parte cuando las condiciones de operación del circuito impliquen variaciones grandes de señal el comportamiento no

lineal se mantendrá, por lo que no se puede modelar al dispositivo por medio de un conjunto de elementos pasivos sino que más bien se tiene que hacer por medio de elementos cuyos efectos guarden una dependencia con respecto a los voltajes de operación aplicados al HEMT.

2. Características de construcción y funcionamiento del HEMT (Maas 1993) y (Singh 1997)

Los HEMT (high electron mobility transistor) son transistores de efecto de campo fabricados en un material de heterojunción. Existen varias maneras de realizar los HEMT; una de tales estructuras posibles es la mostrada en la fig. 1. La parte central de este dispositivo es la interface longitudinal entre una capa de GaAs no dopado y una capa de AlGaAs dopado con silicio. Estas capas son muy delgadas, la capa espaciadora entre ambas es apenas de 50 Årmongs y deben ser fabricadas con técnicas epitaxiales como MBE o MOCVD.

El HEMT opera de un modo semejante al de un MOSFET. El canal es creado en la interfase donde el material de AlGaAs le entrega electrones a la capa de GaAs. Estos electrones son mantenidos en una delgada capa de la superficie. La densidad de electrones en esta delgada capa (llamada gas de electrones bidimensional) es controlada por el voltaje aplicado al *gate*; haciendo el *gate* más positivo se incrementa la densidad de electrones y por lo tanto la corriente del canal también aumenta.

* Ingeniero en Electrónica por la universidad Don Bosco en 1991 y Diplom-Ingenieur Elektrotechnik por la Technische Universität Berlin en 1997. Ha trabajado desde septiembre de 1997 en la Universidad Don Bosco teniendo bajo su responsabilidad, entre otras cosas, las cátedras de Sistemas de Comunicación y Procesamiento Digital de Señales. Sus áreas de investigación están dirigidas a la codificación de señales y técnicas avanzadas de comunicación. E-mail wenceslao.r@hotmail.com



Fig.1: Esquema de la estructura de un HEMT. (2DEG: gas de electrones bidimensional)

Debido a que no hay iones de impurezas en el GaAs que pudieran dispersar a los electrones, ellos se mueven a una alta velocidad a través del material. Como consecuencia, la velocidad de saturación de los electrones y su movilidad son mucho más grandes en el HEMT que en los MESFET: la movilidad a campo eléctrico débil y a temperatura ambiente en el HEMT es de 8000 mientras que en el MESFET es de 5000. Aparte de esto, al reducirse la temperatura la movilidad de los electrones decrece significativamente. Por esta razón, los HEMT a veces son operados a bajas temperaturas reduciéndose con esto la figura de ruido del dispositivo.

3. Características del modelo de gran señal propuesto

3.1 Modelo de gran señal del HEMT

En la operación del HEMT se producen varios efectos, el más importante de ellos es la dependencia de la corriente del canal con respecto al voltaje gate-source. Los otros efectos son por lo general no deseados. Estos son producto de la acumulación y remoción de carga, de las resistencias de trayectoria, del tiempo de retardo de los portadores de carga, etc. Y además estos efectos muestran una dependencia de las tensiones de polarización aplicadas al transistor.

Operando con pequeña señal, en donde los valores de las tensiones mencionadas se apartan muy poco de un valor medio (voltaje de polarización), el HEMT presenta un comportamiento aproximadamente lineal lo cual permite que pueda ser modelado por medio de elementos discretos equivalentes. Los elementos tales como las capacitancias y las resistencias de los contactos, los cuales describen los efectos parásitos externos, al contrario de los elementos internos son independientes del punto de trabajo del HEMT. El circuito equivalente de pequeña señal del HEMT interno utilizado en este trabajo se muestra en la fig. 2. A este circuito se le añaden los elementos parásitos externos para obtener el modelo completo de pequeña señal del HEMT (fig.3)

Para el circuito equivalente de gran señal del transistor existen también varios conceptos, algunos de ellos muy generales. Para el caso que aquí se va a tratar, se ha hecho la simplificación de considerar que el transistor trabaja sólo en las regiones de trabajo acostumbradas y no se dará el caso en que sea polarizado en inversa, por lo cual queda como tarea del usuario tener el cuidado de realizar pruebas únicamente bajo estas condiciones

A continuación se muestra cómo a partir del modelo de pequeña señal se ha desarrollado el modelo de gran señal que aquí se presenta:

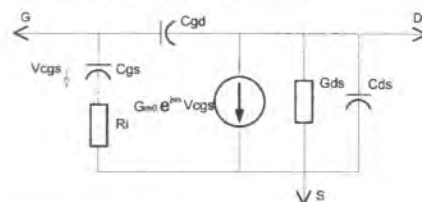


Fig. 2: Circuito equivalente del HEMT intrínseco

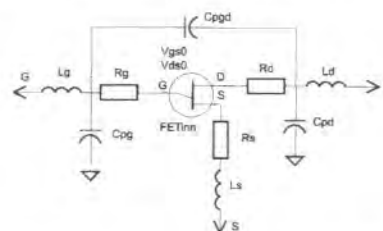


Fig. 3: Modelo de gran señal del HEMT

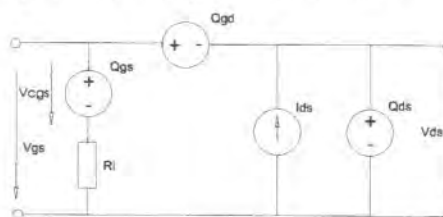


Fig. 4 Transformación directa de los elementos de modelo de pequeña señal al de gran señal.

3.1.2 Representación de los efectos capacitivos

Como un primer paso se sustituyen los elementos del circuito equivalente de pequeña señal del HEMT interno por una combinación de las fuentes de carga y fuentes de corriente (ver Fig.4). Por medio de las fuentes de carga del circuito equivalente de gran señal se busca descubrir los efectos capacitivos del HEMT, los cuales se producen debido a las acumulaciones de carga dentro del transmisor. Estas acumulaciones de carga se acrecientan y remueven siguiendo una determinada relación con respecto a los voltajes en extremos del HEMT. Así es como para tensiones alternas aplicadas al transistor se puede definir que la corriente producida por el movimiento de cargas está dada por la ec.1

$$I = \frac{dQ_{(V_L, V_R)}}{dt} = \frac{\partial Q_{(V_L, V_R)}}{\partial V_L} \cdot \frac{\partial V_L}{\partial t} + \frac{\partial Q_{(V_L, V_R)}}{\partial V_R} \cdot \frac{\partial V_R}{\partial t} \quad (1)$$

En esta ecuación se le designa V_L y V_R a los voltajes de referencia del modelo, en donde V_L es una tensión local, es decir aquella que aparece directamente en extremos del elemento considerado, y V_R es una tensión remota que aparece en alguna otra parte del circuito.

Algo que se ha intentado en este trabajo es romper la discrepancia que existe con respecto a los efectos capacitivos del transistor entre los modelos de gran y de pequeña señal, la cual se hace evidente mediante la siguiente consideración: si un elemento del modelo de gran señal, por ej. una fuerte carga, como arriba se afirmó, muestra una dependencia con respecto a dos magnitudes distintas, significa que la expresión para la variación de la magnitud de ese elemento con respecto al tiempo consta de dos términos (ec.1) en donde se considera que dicha variación (en el caso más general) se debe a la variación con respecto al tiempo de las magnitudes de la cual depende. La discrepancia surge cuando para el caso de operación en pequeña señal, el HEMT tiene un comportamiento tal que el segundo término de la derecha de la ec.1 se vuelve imperceptible y por lo tanto en ese caso el elemento equivalente que suele usarse es una capacitancia que describe únicamente el efecto del primer término de la derecha de la ec.1. El problema que se da como producto de esto, es que usualmente el modelo de gran señal no puede representar al HEMT operando con pequeña señal. El modelo que aquí se presenta puede ser utilizado en ambos modos de operación puesto que se ha aplicado el criterio presentado en (Calvo 1995):

Las fuentes de carga se determinan a partir de las capacitancias del circuito equivalente de pequeña señal (las cuales se dan por conocidas) haciendo uso de la ec.2, en donde Q_0 en principio se puede determinar libremente. Debido a la dependencia con respecto a V_R de la capacitancia de la fuente de carga es así mismo dependiente de V_R , lo cual es correcto desde el punto de vista físico. Para evitar la discrepancia arriba mencionada Q_0 no puede ser cero o constante sino que es igualmente dependiente de V_R y ésta determinada por la ec.3, esto hace que cuando el HEMT opera con pequeña señal los efectos que se producen por las dependencias con respecto a V_R de los dos términos de la derecha de la ec.2, se compensan mutuamente por lo que el efecto neto es una dependencia con respecto a V_L solamente.

$$Q_{(V_L, V_R)} = \int_0^{V_L} C_{(\xi, V_R)} d\xi + Q_{0(V_R)} \quad (2)$$

$$Q_{0(V_R)} = - \int_0^{V_{D0}} C_{(\xi, V_R)} d\xi \quad (3)$$

De acuerdo a lo antes expuesto se tiene para Qgs y Qds:

$$Q_{gs(V_g, V_{ds})} = Q_{gs1} + Q_{gs2}|_{V_g=V_{gs}} = \int_0^{V_g} C_{gs(\xi, V_{ds})} d\xi - \int_0^{V_{gs}} C_{gs(\xi, V_{ds})} d\xi \quad (4)$$

$$Q_{ds(V_g, V_{ds})} = Q_{ds1} + Q_{ds2}|_{V_{ds}=V_{ds0}} = \int_0^{V_{ds}} C_{ds(V_g, \xi)} d\xi - \int_0^{V_{ds0}} C_{ds(V_g, \xi)} d\xi \quad (5)$$

Una dificultad más se da para determinar Qgd puesto que el voltaje local (V_{gd}) en este caso no coincide con ninguno de los voltajes de referencia utilizados (V_g y V_{ds}). Por ello para poder aplicar el mismo método antes expuesto se modificó el circuito equivalente de pequeña señal del HEMT utilizando la llamada transformación de Miller para Cgd (fig 5). De esta manera surgen nuevos elementos conectados de forma tal que la vinculación de V_L y V_R con V_g y V_{ds} es directa.

Las expresiones que definen a los elementos del modelo de gran señal que corresponden a los elementos surgidos de la transformación de Cgd son las siguientes:

- Para la capacitancia *gate - source*:

$$Q_{gd_gs(V_g, V_{ds}, V_{ds0})} = \int_0^{V_g} C_{gd(\xi, V_{ds})} d\xi - \int_0^{V_{gs}} C_{gd(\xi, V_{ds})} d\xi \quad (6)$$

- Para la fuente de corriente del *puerto gate source*:

$$-Q_{gd_ds(V_g, V_{ds}, V_{ds0})} = - \int_0^{V_{ds}} C_{gd(V_g, \xi)} d\xi + \int_0^{V_{ds0}} C_{gd(V_g, \xi)} d\xi \quad (7)$$

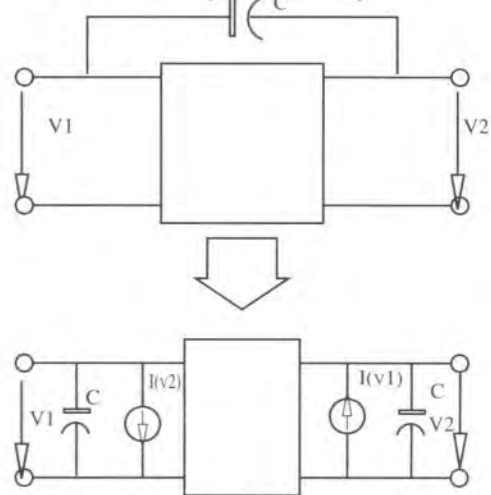


Fig. 5 Transformación de Miller para Cgd

- Para la capacitancia *gate - source*:

$$Q_{gd_gs(V_g, V_{ds}, V_{ds0})} = \int_0^{V_g} C_{gd(\xi, V_{ds})} d\xi - \int_0^{V_{gs}} C_{gd(\xi, V_{ds})} d\xi \quad (6)$$

- Para la fuente de corriente del *puerto gate source*:

$$-Q_{gd_ds(V_g, V_{ds}, V_{ds0})} = - \int_0^{V_{ds}} C_{gd(V_g, \xi)} d\xi + \int_0^{V_{ds0}} C_{gd(V_g, \xi)} d\xi \quad (7)$$

Con los cambios anteriormente mencionados el circuito equivalente de gran señal puede representarse según la Fig.6.

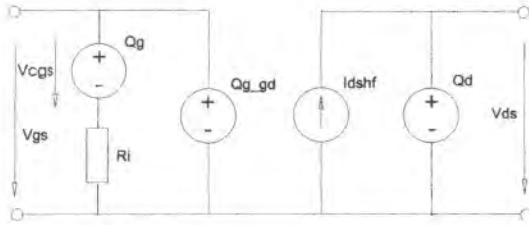


Fig. 6. Modelo de gran señal del HEMT intrínseco.

Las ecuaciones que definen a las fuentes de carga del circuito de la Fig. 6. en función de las fuentes de carga antes definidas son:

$$Q_g = Q_{gs} = Q_{gs1} + Q_{gs2} \quad (10)$$

$$Q_{g_gd} = Q_{gd_gs} - Q_{gd_ds} \quad (11)$$

$$Q_d = Q_{ds} - Q_{g_gd} = Q_{ds1} + Q_{ds2} - Q_{gd_gs} + Q_{gd_ds} \quad (12)$$

3.1.3 Representación del control de la corriente del canal

En el circuito equivalente de gran señal la corriente de *drain* es suministrada por una fuente de corriente (Fig. 4) y depende de los voltajes V_{gst} y V_{ds} . V_{dcgst} es el voltaje que aparece en extremos de Q_{gs} pero con τ segundos de retraso.

Para un punto de trabajo fijo y para pequeña señal se permite que las contribuciones de V_{gs} y V_{ds} a la corriente de *drain* se puedan tratar por separado por medio de una fuente de corriente que depende de V_{dcgst} y una resistencia conectada entre *drain* y *source* (Fig. 2).

De las fuentes bibliográficas¹ se conocen dos métodos por los cuales se puede determinar una expresión para la corriente del canal.

Uno de estos métodos utiliza una conexión en paralelo de una fuente de corriente con un circuito serie RC (Fig. 7). La fuente de corriente proporciona los valores que corresponden a la curva característica de voltaje - corriente de corriente directa del HEMT y el circuito RC toma en cuenta los efectos dispersivos que aparecen cuando se trabaja con alta frecuencias y que se manifiestan como un incremento de la admitancia de salida (Felgentreff 1995) (Fig. 8).

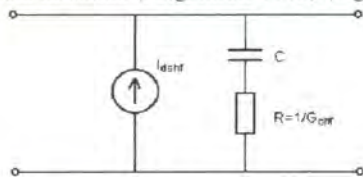


Fig. 7: Circuito equivalente para el modelado de la corriente de *drain* (primera alternativa)

1p.ej.de Felgentreff

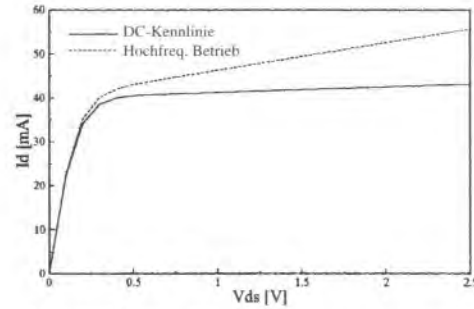


Fig. 8: Representación a modo de ejemplo del comportamiento de la corriente de *drain* con respecto a V_{ds} para DC y para alta frecuencia. La línea continua es la curva de DC y la línea cortada es la curva para alta frecuencia.

Por el segundo método, la corriente se determina por medio de la integración, de g_m y de g_{ds} (ec. 13). Además se demostró en (Koppel 1996) que esta integración debe ser independiente del camino de integración que se utilice, es decir, que por medio de la elección adecuada de éste, la ecuación 13 puede simplificarse tal y como, por ejemplo se muestra en las ecuaciones 14 y 15, donde la integración se lleva a cabo sólo con uno de los parámetros. Además, sucede que debido a posibles errores de medición no se puede descartar que existan algunas diferencias entre los resultados de las ecuaciones 14 y 15. En este caso se recomienda en (Felgentreff 1995) utilizar la ecuación donde se trabaja con g_m (ec. 15), puesto que este parámetro se puede determinar de las mediciones con mayor exactitud.

$$I_{dshf}(V_{gs}, V_{ds}) = \int_{V_{gs}=\min}^{V_{gs}} g_m(\xi, V_{ds}) d\xi \bigg|_{V_{ds}=\min}^{V_{ds}} + \int_{V_{ds}=\min}^{V_{ds}} g_{ds}(V_{gs}, \xi) d\xi \bigg|_{V_{gs}=\min}^{V_{gs}} \quad (13)$$

$$I_{dshf}(V_{gs}, V_{ds}) = \int_{V_{gs}=\min}^{V_{gs}} g_m(\xi, V_{ds}) d\xi \bigg|_{V_{ds}=0}^{V_{ds}} + \int_{V_{ds}=0}^{V_{ds}} g_{ds}(V_{gs}, \xi) d\xi \bigg|_{V_{gs}=\min}^{V_{gs}} \quad (14)$$

$$= \int_{V_{gs}=\min}^{V_{gs}} g_{ds}(V_{gs}, \xi) d\xi \bigg|_{V_{ds}=\min}^{V_{ds}} \quad (15)$$

$$I_{dshf}(V_{gs}, V_{ds}) = \int_{V_{gs}=\min}^{V_{gs}} g_m(\xi, V_{ds}) d\xi \bigg|_{V_{ds}=\min}^{V_{ds}} + \int_{V_{ds}=\min}^{V_{ds}} g_{ds}(V_{gs}, \xi) d\xi \bigg|_{V_{gs}=\min}^{V_{gs}} \quad (15)$$

$$= \int_{V_{gs}=\min}^{V_{gs}} g_m(\xi, V_{ds}) d\xi \bigg|_{V_{ds}=\min}^{V_{ds}} \quad (15)$$

Por medio de unas pruebas en las que se buscaba determinar cual era el mejor método se obtuvieron los siguientes resultados:

a) El primer método se ha descartado porque tanto los valores de la transconductancia y de la impedancia de salida que se obtienen de la curva DC de I vrs. V no coinciden con los valores de g_m y g_{ds} extraídos para el circuito equivalente de pequeña señal. Además ha existido el problema de que en el

programa de simulación utilizado, la implementación de esta admitancia de salida no ha sido posible.

b) Para la aplicación del segundo método se observó que los resultados obtenidos al utilizar las ec. 14 y 15 eran distintos entre sí. Como solución se optó por utilizar una segunda fuente de corriente, que depende únicamente de V_{ds} , que compense el comportamiento de la primera (que es determinada por la ecuación 15) con respecto a la admitancia de salida.

$$I_{ds}(V_{gs}, V_{ds}) = I_{dshf(V_{gs}, V_{ds})} + I_{gds(V_{gs}, V_{ds})} \Big|_{V_{gs}=V_{gs0}} \quad (16)$$

$$I_{gds(V_{gs}, V_{ds})} \Big|_{V_{gs}=V_{gs0}} = \int_0^{V_{ds}} (g_{ds}(V_{gs}, \varepsilon) - \frac{\partial dshf(V_{gs}, V_{ds})}{\partial V_{ds}}) d\varepsilon \Big|_{V_{gs}=V_{gs0}} \quad (17)$$

Donde V_{gs0} es el voltaje retrasado τ segundos del voltaje que aparece en extremos de Cgs. V_{gs1} se igualará a V_{gs0} (constante) para cuando se opere con pequeña señal y se igualará a V_{gs} (variable) para cuando se opere con gran señal; de esta manera se podrá simular el funcionamiento correcto del transistor en ambos modos de operación.

Para la operación en gran señal se incurre en el error de que al hacerse V_{gs1} igual a V_{gs} la ec.16 se reduce a la ec. 14 de donde la transconductancia efectiva no coincide con los valores de pequeña señal.

4. Modelo de tablas

La forma corriente que se tiene para efectuar el modelado de gran señal de un transistor es determinar una expresión cerrada para cada uno de los elementos del circuito equivalente en donde las variables de las cuales dependen las magnitudes de estos elementos son los voltajes que se aplican al transistor, en este caso V_{gs} y V_{ds} . Sin embargo, existe una forma alternativa para definir esta dependencia y es a través de una tabla de valores, lo que se tiene entonces son valores de carga, corriente, etc. (según el elemento del que se trate), donde cada valor es aquel que el elemento tendría para un determinado valor de V_{gs} y de V_{ds} . Se pretende tener una tabla(Matriz) que contenga un número de valores que correspondan a un número suficiente de puntos de trabajo (combinaciones de V_{gs} y V_{ds}). Este método exige, por supuesto, que el programa de simulación tenga la capacidad de efectuar interpolaciones para determinar valores que no coincidan con puntos de operación almacenados en la matriz (puntos intermedios).

Las ventajas de usar este método son las siguientes:

- Se evita tener que determinar una expresión que siendo fácil de diferenciar cubra con todos los valores obtenidos de la medición.

- Los valores medidos pueden ser determinados con exactitud (al buscar en la matriz).
- El proceso de interpolación del programa de simulación permite que una diferenciación pueda hacerse sin mayor dificultad.

5. Procedimiento para la determinación de los parámetros del modelo

Los pasos seguidos para la creación de las tablas son:

- Extracción de los parámetros del circuito equivalente de pequeña señal del HEMT para un número adecuado de puntos de trabajo (de acuerdo al concepto de la fig. 2).
- Conversión de los valores de los elementos del transistor intrínseco de modo que estos valores correspondan a aquellas condiciones de trabajo para las cuales las combinaciones de V_{gs} y V_{ds} (que aparecen en el interior del HEMT) formen una matriz ortogonal de valores.
- Resolución numérica de las integraciones para el cálculo de los elementos de las tablas para las fuentes de carga y corriente.
- Creación de un archivo de formato CITIfile para almacenar los resultados.
- Creación de una Dataset para el MDS a partir del archivo con los resultados.

6. Comprobación del funcionamiento del modelo

A continuación se muestra en qué medida el modelo de gran señal del HEMT aquí presentado se ha aproximado a la meta deseada.

Análisis de gran señal

El modelo desarrollado se ha utilizado para simular el funcionamiento de un HEMT en un mezclador que debe funcionar en la banda X, que corresponde al rango de 8 a 12 Ghz, (RF), y que debe producir una FI de 400 Mhz. El diagrama esquemático del circuito se muestra en la Fig.9. Los resultados de la simulación se han comparado con valores medidos de ganancia y ambos se presentan en la gráfica de la fig. 10.

Los resultados pueden interpretarse de la siguiente manera: por una parte se observa que la curva que corresponde a la simulación presenta la misma ten

dencia que evidencian los valores medidos; sin embargo, para bajas frecuencias y para el punto de máxima ganancia los resultados de la simulación se apartan de los de la medición, lo cual conduce a pensar en posibles errores que habrían en la representación de la amplificación de las impedancias de entrada o salida por parte del modelo: Una causa de error ajeno al modelo del HEMT es que en el circuito se ha trabajado con cierta imprecisión debido a que las inductancias en espiral no han podido ser modeladas adecuadamente, lo que afecta al comportamiento en frecuencia del circuito.

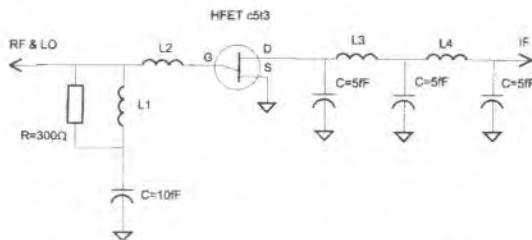


Fig. 9: Diagrama esquemático del mezclador utilizado para probar el modelo desarrollado.

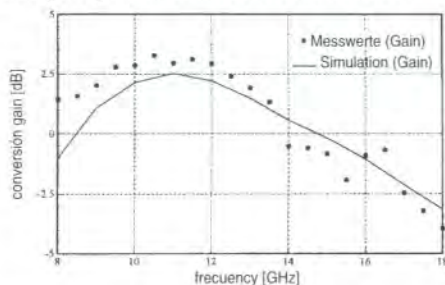
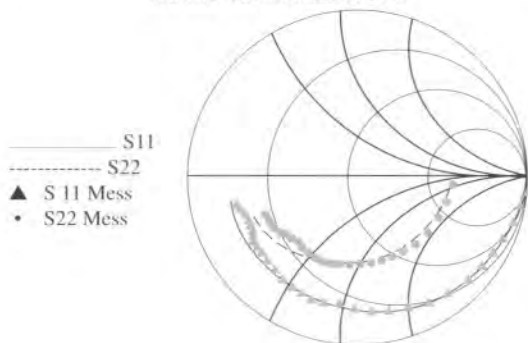


Fig.10: Ganancia de conversión del mezclador. La línea continua es el resultado de la simulación y los puntos son valores medidos.

Análisis de pequeña señal

La forma inmediata para comprobar el funcionamiento del modelo operando con pequeña señal, es a través del análisis de parámetros distribuidos (Fig.11.)

S12 y S21 en la región activa:
S12 und S21 in Anlaufbereich



S12 y S22 en la región activa:
S12 und S21 in Anlaufbereich

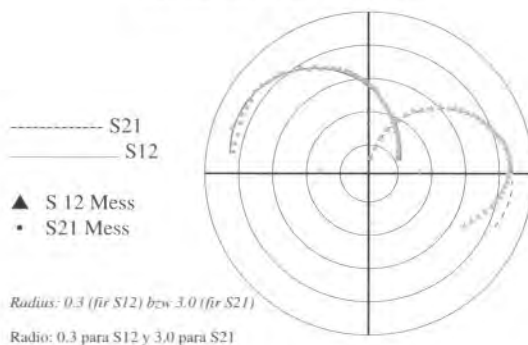


Fig. 11: Comparación de parámetros S. Intervalo de frecuencias: 0.5 -40 GHz. $V_{gs}=-0.2$ V, $V_{ds}=0.4$ V.

(Las líneas corresponden a la simulación y los puntos a los valores medidos)

7. Conclusión

En este documento se han mostrado las virtudes y las deficiencias del desarrollo y empleo de un modelo de gran señal apoyado en tablas de datos. Se ha demostrado que este modelo es una herramienta muy práctica que puede servir para el diseño de circuitos no lineales contando sin embargo con algunos inconvenientes tales como que para ciertas regiones en la creación de las tablas surgen errores que conducen a resultados equivocados (Rivas 1997). Otro inconveniente del modelo es que los valores de voltaje para el punto de trabajo del HEMT deben ser colocados manualmente por el usuario. Se tiene así mismo que sobre la operación en gran señal aún son necesarias otras mediciones y análisis para determinar la causa de los errores que se han detectado.

Referencias

- Calvo, M.V; Snider, A. D. y Dunleavy, L.P. (1995). "Resolving capacitor discrepancies between large and small signal FET models". TH3B-6. MTT-5. 1251-1254.
- Felgentreff, T. (1995). Tesis doctoral "Großsignal - und Rauschmodellierung von Heterostruktur-Feldeffekttransistoren". Univ. Téc.de Munich.
- Koppel, T. (1996). Trabajo de graduación "Entwicklung eines Großsignalmodells Für GaAs MESFETs" Univ. Téc. de Berlín.
- Mass, Stephen A. (1993). Microwave Mixers. Artech House. Londres.
- Petermann, K. (1994). Hojas de trabajo para la cátedra "Hochfrequenztechnik II" TU Berlín
- Rivas Z., O.W. (1997). Trabajo de graduación "HEMT Modellierung". Univ. Téc.de Berlín
- Singh, J. (1997). Dispositivos Semiconductores. McGraw-HILL. México.

Introducción a la técnica P.I.V. para la medición de flujo

Oscar G. Durán Vizcarra*

Resumen

La finalidad de este artículo es presentar la técnica PIV (Particle Image Velocimetry) como una alternativa a las necesidades de medición de flujo en condiciones especiales, ya que los sistemas de medición de flujo más conocidos no permiten obtener tanta información como ella. En aras de esto, se traza un bosquejo de la evolución en las técnicas avanzadas de medición de flujo hasta llegar al caso que nos ocupa. Posteriormente se procede a describir un sistema PIV genérico.

1. Evolución de las técnicas de mapeo

El caso más simple de obtención de información de un flujo lo constituye la vieja técnica de humedecerse la punta de uno de los dedos y alzarlo para determinar la dirección del viento al sentir que se enfría el lado del dedo sobre el cual choca el frente de aire. Evidentemente no se puede obtener más que información superficial por estos medios. Sin embargo, la técnica de medición de flujo “*de alambre caliente*” está basada en este principio. El efecto de enfriamiento de un flujo a través de uno o más alambres o películas calientes brinda una lectura del flujo que puede ser calibrada a un valor absoluto de velocidad.

A lo largo del tiempo han ido apareciendo técnicas de medición de flujo cada vez más sofisticadas, así tenemos el caso de técnicas como LDA¹, visualización de flujo propiamente dicha, mapeo de flujo, y PIV entre otras. Cada una de estas técnicas presenta una serie de atributos y desventajas que las hacen aplicables en determinadas circunstancias.

En el caso de LDA, su principal ventaja consiste básicamente en la remoción de la necesidad de calibración, así como de introducir hardware al flujo. Esto hace de LDA una técnica altamente exacta y no intrusiva [6]. Los sistemas LDA comerciales son sistemas totalmente integrados con interfaces y software para el control computarizado del hardware de adquisición de datos. La información de velocidad puede ser muestreada a varios megahertz para suministrar información con un extremadamente alto grado de resolución acerca de la evolución en el tiempo del flujo en un punto. La información acerca de la distribución espacial de velocidades se obtiene desplazando el volumen de medición de un punto de

prueba a otro; muestreando la velocidad en cada punto hasta construir un mapa de velocidades. Esta propuesta ve limitada su utilidad cuando el flujo no es completamente estable o repetible y el resultado de las mediciones es obtener una imagen de promediada del campo de flujo. Información sobre transientes, inestabilidades y fenómenos aleatorios se pierde debido a esto, revelando solo características del flujo independientes del tiempo. El posicionamiento de dos o más puntos de prueba en el flujo en distintos lugares permite el cálculo de correlaciones espaciales, lo que brinda información estadística acerca de las escalas de tiempo y espacio presentes en el flujo.

Por otra parte, en el pasado la información acerca de la estructura del flujo ha sido obtenida por técnicas de mapeo más simples. Interpretada junto con los resultados de mediciones puntuales, la visualización de flujo puede suministrar algún detalle de flujos complejos y no estacionarios, pero presenta algunas limitaciones obvias: la información en visualización de flujo es esencialmente cualitativa y aunque algunos puntos en la imagen pueden tener valores cuantitativos tentativamente asociados, la información espacial que se deriva de tal imagen es de bajo orden; las dimensiones de las capas, las áreas de vórtice y recirculación pueden ser medidas manualmente, pero información de mayor orden como la velocidad, gradientes de velocidad, correlación espacial, vorticidad, etc. no se logra a niveles de certeza útiles.

El mapeo cuantitativo de flujo es una extensión obvia de las técnicas de visualización. El pariente más obvio de PIV es la técnica *Streak Velocimetry* (velocimetría de trazos); si uno fuera a medir la longitud de los trazos registrados durante un período de exposición conocido podría determinar la velocidad en el punto. Midiendo la longitud de todos los trazos en la imagen se puede obtener información de las velocidades sobre un área extendida de flujo.

No obstante, hay algunos problemas: la longitud de los trazos no es la distancia que ha viajado la partícula; el trazo es formado por una partícula

* Ingeniero en Electrónica por la Universidad Don Bosco, y ha realizado cursos de Doctorado en Ingeniería Mecatrónica de la Universidad Carlos III de Madrid. En la actualidad se desempeña como Director de la Escuela de Ingeniería Electrónica de la Universidad Don Bosco. E-Mail: oscardu@vianet.com.sv

¹ LDA: Laser Doppler Anemometry

trazadora que se mueve en el material de registro durante cierto tiempo de exposición extendido y esta imagen no es infinitamente pequeña, así que el trazo es alargado al principio y al final por las dimensiones del trazador. Para determinar la distancia recorrida por la partícula debe analizarse el trazo en la imagen a fin de corregir el tamaño finito de la partícula, el efecto de ensanchamiento de la lente y posiblemente la respuesta no lineal del medio de registro. Las partículas trazadoras son hechas visibles por una lámina de luz. Dificultades adicionales son encontradas por partículas saliendo de la lámina de luz, lo que produce trazos cortados. Los trazos que se cruzan, además, suponen dificultades para el análisis cuantitativo de la imagen. El empleo de análisis de imagen para la evaluación cuantitativa de fotografías de trazos ilustra el ímpetu del desarrollo del mapeo cuantitativo de flujo.

2. Fundamentos de PIV

PIV es un método óptico para efectuar diagnósticos de flujo que ya es aplicado en importantes problemas. En PIV se mide la velocidad instantánea del flujo simultáneamente en muchos puntos de una sección transversal del mismo, permitiendo la determinación de derivadas y correlaciones espaciales de la velocidad [4].

Básicamente la técnica consiste en que conocido un tiempo de muestreo se miden los desplazamientos en imágenes consecutivas de las partículas inyectadas en el flujo, a partir de lo que se determinan las velocidades. La capacidad de suministrar mediciones de alta resolución de gradientes de velocidad turbulentos es única y propia de esta técnica y es una de las razones por las cuales goza en la actualidad de especial atención.

Aunque la mayoría de aplicaciones de PIV han hecho gala de sus bondades, el trabajo continuo necesita investigar su validez en mediciones turbulentas y mejorar su operación. También es imperativo resolver algunos problemas o limitaciones tecnológicas que en el presente obstaculizan aplicaciones más amplias del método.

Cabe mencionar a este respecto los elevados costes, el gran tamaño y la escasa eficiencia de las fuentes de luz láser utilizadas para iluminar las partículas en la medición de la sección de flujo. Otro problema tecnológico concierne al registro de las imágenes PIV, donde la elección de hoy día se encuentra entre la alta resolución, lentitud y lo engorroso del proceso químico fotográfico y la rapidez y menor resolución de los detectores matriciales electrónicos.

Hay también que tener en cuenta otras inconveniencias o desventajas tanto de carácter tecnológico como metodológico es la relativa lentitud del procesamiento de las imágenes PIV (o el relativamente elevado coste de obtener un procesamiento más poderoso).

El procesamiento óptico de imágenes PIV está basado principalmente en el análisis de Fourier. Una transformada de Fourier puede trabajar instantáneamente sobre una imagen bidimensional y es fácilmente implementada en la práctica, así que cada vez que se necesita un gran número de transformadas de Fourier en tareas de procesamiento, siempre hay potencial disponible para aumentar el trabajo sin requerir invertir mayores recursos utilizando métodos ópticos.

Tal como es a menudo el caso, en el procesamiento óptico, el cuello de botella aparece al momento de escribir los datos leídos y luego presentar los resultados. El procesamiento de datos podría ser ejecutado completamente por métodos numéricos. Sin embargo, los datos, que a menudo están en forma de imagen fotográfica, aun deben ser convertidos a un formato electrónico digital y luego transferidos a la computadora. Incluso cuando las imágenes de partículas son directamente capturadas por una cámara de vídeo, la transferencia de las imágenes toma una cierta cantidad de tiempo. No obstante, una vez que este proceso ha sido culminado el resto del procesamiento es materia de cálculos digitales y la velocidad de procesamiento es determinada por la velocidad y organización del procesador digital.

Tanto los componentes de procesamiento óptico como las computadoras digitales mejoran continuamente y el coste, en particular del poder de procesamiento digital, continúa en caída, de manera que la evaluación de la ejecución relativa debe ser frecuentemente actualizada.

Dada la madurez que están alcanzando las técnicas de medición PIV, podríamos ver que hay una necesidad emergente de procesadores electrónicos dedicados y optimizados para los requerimientos particulares del procesamiento de grabaciones PIV. Tales procesadores habrán eventualmente de sobrepasar los que hasta la fecha se han ensamblado tanto para propósitos generales como para procesamiento óptico, en términos de velocidad y exactitud.

A continuación se procede a describir de forma breve las principales etapas que constituyen un sistema de PIV genérico.

3. Alimentación de partículas

El rango de materiales con que se alimenta el flujo es muy amplio, así se puede tener partículas del orden de las micras en túneles de viento a velocidades

supersónicas que bien podrían ser astillas de madera, o incluso pequeños discos de papel inyectados en flujos lentos [4]. En general no hay un límite superior en cuanto al tamaño de las partículas, pudiéndose aplicar la técnica a medir velocidad en desplazamiento de carros en una autopista, ovejas en una colina o nubes en el cielo.

Las aplicaciones de mayor demanda son regímenes de alta velocidad, de decenas de m/s en adelante. En estos casos las partículas alimentadas han satisfecho el conflicto de requerimientos que supone dispersar suficiente luz mientras son lo bastante pequeñas para seguir el flujo.

Hay requerimientos auxiliares como que el material a inyectar debe ser barato para poder ser empleado en cantidades aceptables dado el volumen de flujo en cuestión y no ser contaminantes ni peligrosamente inestables.

Con el objeto de que las partículas trazadoras no impongan al flujo perturbaciones que hagan poco fiables las lecturas se suele hacer consideraciones [4] como las de las siguientes características:

3.1 Efectos inerciales

Diferencias de densidad grandes entre el fluido y las partículas provocan retrasos en la respuesta de las partículas a la aceleración del fluido. Esto es significativo cuando hay fluctuaciones de alta frecuencia en el flujo.

3.2 Efectos gravitatorios

Grandes diferencias de densidad entre las partículas y el fluido además pueden inducir cierto grado de sedimentación de las primeras tendiendo a ubicarlas en capas superiores o inferiores, según sea la densidad relativa de las partículas al fluido. Esto podría resultar insignificante si la velocidad del flujo es tal que mantiene homogeneidad en la distribución de las partículas a lo largo del flujo.

3.3 Promediación espacial de la turbulencia

Si en el flujo se presentan remolinos y las partículas son de mayor tamaño que estos puede perderse información acerca de movimientos de pequeña escala en el flujo. Esto no es significativo en la mayoría de estudios.

4. Registro

La configuración genérica del sistema PIV mostrado en la Fig. 1 indica que el haz de luz laminar ilumina a las partículas de forma que el arreglo de ópticas o lentes

está dispuesto en un plano paralelo al de dicho haz. La exposición sobre la superficie de registro debe ser lo bastante buena como para producir alto contraste a fin de no dificultar la identificación de las partículas sobre el fondo. La exposición es el producto de la intensidad de la iluminación y el tiempo que esta incide. La duración es relativamente corta², así que la intensidad de la iluminación debe ser elevada.

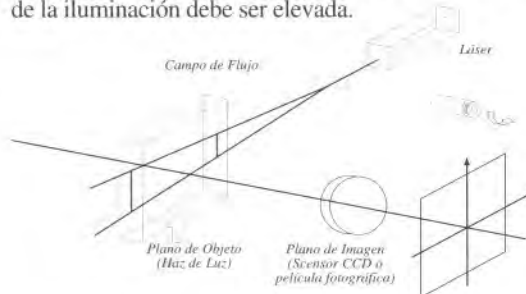


Fig. 1: Esquema de la captación de las imágenes PIV.

La luz dispersada por las partículas inyectadas al flujo es colectada por un arreglo de lentes a fin de formar una imagen en el material de registro.

Estudios acerca de la potencia dispersada por partículas dentro de un rango de diámetros que van de 0.2 a 15 μm indica que esta potencia tiende a aumentar rápidamente para partículas con diámetros mayores o iguales a 1 μm y más lentamente abajo de este valor.

La posición de las partículas puede ser registrada fotográfica o electrónicamente [1][4]. El registro electrónico es más práctico y conveniente debido a que la imagen es directamente convertida a formato digital; sin embargo, la resolución de los detectores matriciales disponibles había sido inferior que la de las películas fotográficas y generalmente muy baja para estudios detallados de flujo turbulento.

Una evaluación más detallada y una comparación entre ambos métodos habría de tomar en cuenta otros parámetros tales como la sensibilidad a la luz, el rango dinámico de la imagen registrada y el ruido de fondo. No obstante, la tendencia actual es desplazarse hacia mecanismos de registro de estado sólido³ y de hecho sistemas con resolución de 2K • 2K ya no resultan poco frecuentes.

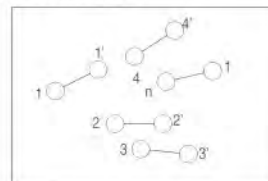


Fig. 2: Desplazamientos de las partículas entre imágenes consecutivas ($\Delta_1, \Delta_2, \Delta_3, \Delta_4, \dots, \Delta_n$).

² Típicamente alrededor de 10 ns

³ CCD: Charge-Coupled Device

Para nuestros propósitos basta decir que un registro PIV consiste en una doble exposición de una multitud de partículas en el flujo. Cuáles imágenes de partículas son de hecho capturadas en el registro es determinado por la sección de flujo, la cual es iluminada por el rayo de luz y se encuentra dentro del campo visual del detector.

La luz es concentrada en una delgada “lámina” de luz para iluminar una sección de cruce del flujo. La región iluminada y el campo visual del detector definen la región de medición.

El haz laminar puede ser generado ya sea por medio de un arreglo de lentes o por medio de la rotación y reflexión de un haz láser normal sobre espejos [1][4].

En el primero de los casos una lente cilíndrica frente a la fuente láser se encarga de dispersar la luz en una sola dirección formando la lámina en cuestión. Agregando otra lente cilíndrica y dos o más esféricas se consigue el ajuste necesario del ángulo y estrechez de la lámina generada tal como se ilustra en la Fig. 3.

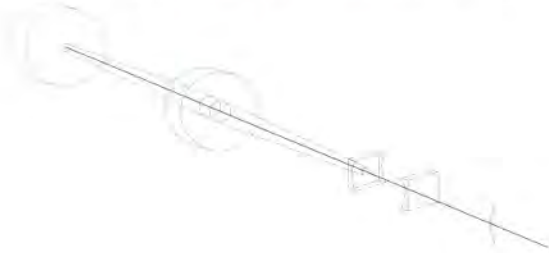


Fig. 3: Generación del haz laminar por medio de lentes.

La otra técnica de generación del haz laminar típicamente empleada consiste en buscar la asistencia de arreglos mecánicos o electroópticos que modulen la posición del rayo originalmente obtenido a partir de la fuente. Esta es una técnica ampliamente utilizada, pese a que en condiciones extremas tiende a limitar la cantidad de luz que atraviesa el flujo. De hecho, la razón entre la duración del pulso de luz y la separación entre pulsos suele ser de solo 10, es decir que solo un 10% de la luz es empleada realmente.

Un método más efectivo para la formación de los haces pulsantes consiste en generar pseudo-láminas mediante el uso de un espejo parabólico (ver Fig. 4). El láser es primero adelgazado usando dos lentes esféricas y luego dirigido a un espejo poligonal rotatorio que tiene típicamente 8, 12 ó 16 lados. El abanico de luz así producido es entonces reflejado sobre el espejo parabólico colimador. De esta manera la eficiencia del sistema de iluminación puede ser incrementada considerablemente. Es de hacer notar también que una

lámina láser generada de esta forma posee una intensidad muy uniforme comparada con el perfil Gaussiano que corresponde a la de las producida mediante lentes cilíndricas

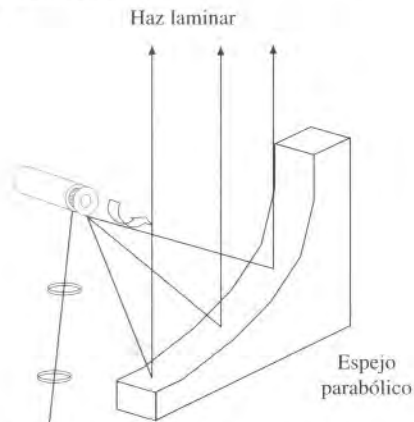


Fig. 4: Generación del haz laminar por medio de reflexión.

El volumen del flujo determinado por el rayo de luz utilizado y la proyección del área de interrogación define el volumen de medición. Un parámetro de importancia a considerar es el número de pares de imágenes de partículas dentro de un área de interrogación, a menudo llamado Densidad de Imagen.

Como se asume que las partículas están aleatoriamente distribuidas en el flujo con una densidad de probabilidad uniforme, el número de imágenes de partículas registradas obedecerá a un proceso de Poisson y la distribución estadística del número de pares es descrita por la estadística de Poisson. Con el fin de obtener información óptima en los registros PIV es usual inyectar partículas con densidades de imagen entre siete y veinte⁴. En general la técnica PIV suele manejar un volumen del orden de 10^9 a 10^{11} partículas/s.

5. Procesamiento

Las tareas de procesamiento de imágenes PIV se dividen en dos partes: detección de partículas y establecimiento de pares. Es necesario identificar la posición de las partículas en las parejas de imágenes y para ello se parte de lo siguiente:

- Las partículas tienen mayor intensidad que su fondo.
- Las partículas son típicamente mayores que la menor resolución del medio de registro (1 pixel si se trata de CCD).

⁴ La densidad de imagen es una cantidad adimensional

- La intensidad del fondo suele no ser uniforme.

La alternativa más simple para identificar las partículas se basa en combinar procedimientos de umbralizado y binarización [3][10] (ver Fig. 5). Esto consiste en reasignar valores a la intensidad de cada pixel, tomándose como 1 aquellos que excedan cierto valor (umbral) y como 0 aquellos que no lo excedan. Este procedimiento tiene como principal ventaja su sencillez y rapidez de operación, pero a veces resulta problemático en flujos en los que las partículas tienden a aglomerarse en zonas determinadas produciendo grandes concentraciones de ellas.

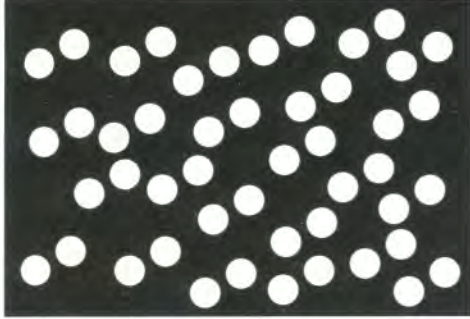


Fig. 5: Umbralización aplicada a la identificación de las partículas trazadoras respecto del fondo.

Otra alternativa consiste en aplicar algoritmos de detección de bordes [3][10] (ver Fig. 6), que básicamente consisten en detectar la frontera entre regiones con distintos niveles de gris en una imagen. Las líneas obtenidas por este procedimiento proporcionan el contorno de las partículas uno de los problemas que presenta es que computacionalmente requiere mayor tiempo de procesamiento.

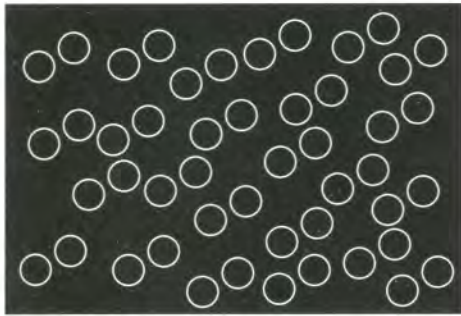


Fig. 6: Detección de bordes aplicada a la identificación de las partículas trazadoras respecto del fondo.

También puede recurrirse a algoritmos de correlación [3]. Esta opción disminuye la influencia

de irregularidades en fondo que contenga la imagen sobre la obtención de características de la misma.

Por otra parte, información adicional acerca de la posición real de las partículas puede obtenerse si se conoce bien el aspecto de las partículas. Esta es una alternativa extremadamente buena, si bien requiere un cierto conocimiento previo de las características de las partículas.

Una vez que la correlación ha sido llevada a cabo la posición de las partículas se determina por detección de máximos (ver Fig. 3) y el posicionamiento de estos sobre el plano.

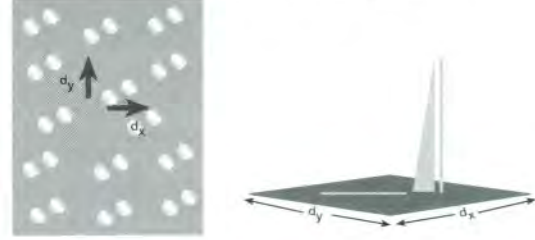


Fig. 5: Detección de máximos en la función de correlación.

Cuando ya se tienen las doble exposiciones de las imágenes de partículas, el siguiente paso consiste en determinar los desplazamientos de las partículas. Como se ha explicado anteriormente, si para identificar las partículas sobre el fondo se recurrió al método de correlación, como resultado de esta ya casi se tienen determinados los desplazamientos, en caso de haber recurrido a una de las otras técnicas habrá de ejecutarse un algoritmo específico para tales tareas.

Es necesario evaluar la separación y dirección de los bordes de las partículas, siendo las componentes de esa separación $\mathbf{d} = d_x, d_y$; esta información permite deducir las componentes del desplazamiento $\Delta = \Delta_x, \Delta_y$ y conociendo el índice de magnificación M^5 y el tiempo Δt entre exposiciones, la velocidad del flujo tiene las componentes $\Delta = \Delta_x, \Delta_y$.

Las expresiones que relacionan las variables antes mencionadas son las siguientes:

$$d_x = \frac{\lambda f}{\Delta x} \quad \text{y} \quad d_y = \frac{\lambda f}{\Delta y} \quad (1)$$

$$\Delta_x = \frac{\Delta x}{M} \quad \text{y} \quad \Delta_y = \frac{\Delta y}{M} \quad (2)$$

$$u_x = \frac{\Delta x}{\Delta t} \quad \text{y} \quad u_y = \frac{\Delta y}{\Delta t} \quad (3)$$

donde λ es la longitud de onda del láser y f es la distancia focal del sistema óptico.

⁵ Donde M se define como la relación entre las dimensiones de las imágenes original y captada (Ver Fig. 6)

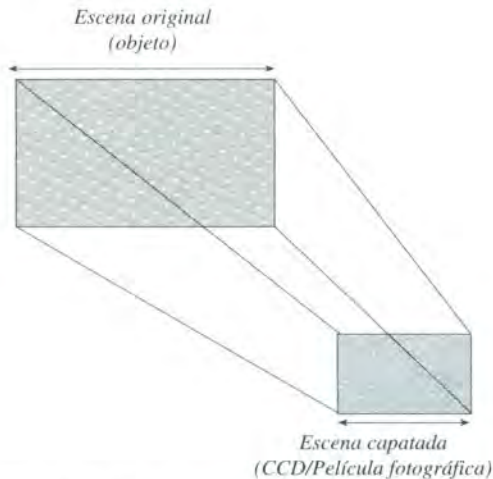


Fig. 6: Relación entre la imagen original y la captada.

El siguiente paso es recoger el patrón arrojado por las imágenes y obtener una transformada bidimensional de Fourier, para luego determinar el módulo o el cuadrado de la misma para fines de cómputo. También suele hacerse estimaciones de la relación señal-ruido de la medición a fin de mejorar su calidad.

Cuando las posiciones de las partículas ya han sido determinadas, cualquiera que haya sido el algoritmo empleado, es menester proceder a establecer los pares de ellas, a partir de lo cual se puede calcular los desplazamientos que toman lugar. Para ello se utilizan complejos algoritmos [2][7][9] que van desde técnicas puramente estadísticas hasta una combinación de ellas con mecanismos de inteligencia artificial.

Dado que prácticamente todo el volumen de información de estos procesos de medición es llevado a cabo por computador, resulta relativamente fácil dar una presentación adecuada y moderna a los resultados. En la actualidad ya se cuenta con sistemas de medición PIV comerciales completos tanto en términos de hardware como de software.

6. Aplicaciones

Las aplicaciones de la técnica se encuentran presentes en aquellos casos en los que se requiera efectuar mediciones en flujos turbulentos o donde se requiera efectuar mediciones instantáneas.

Ya que PIV es una técnica de medición de campo completo, en oposición a aquellas técnicas en las que se busca un punto representativo que presente las características promedio del campo de flujo, es posible obtener de ella mucha información adicional.

Lo expresado anteriormente nos conduce a casos como los que se citan en los ejemplos presentados a continuación:

6.1 Simulación de microrráfagas atmosféricas [8]

Se simula y describe el efecto de masas de aire a través de la atmósfera que tienden a caer sobre la tierra. Este fenómeno aparece a veces junto con tormentas fuertes y, debido a las variaciones en la componente horizontal de la velocidad del viento asociadas a él puede suponer serios problemas para aeronaves que vuelan a baja altura.

6.2 Flujo en las vecindades de una válvula prostética [5]

Un requisito esencial de las válvulas prostéticas es que no han de producir altas tasas de vorticidad en los flujos de sangre en sus proximidades, debido a que ello incide directamente sobre la hemólisis. Como sabemos, a diferencia de la mayoría de técnicas de medición de flujo, PIV permite efectuar lecturas en flujos con vorticidad, lo que lo hace idóneo para esta aplicación.

6.3 Turbulencia en rejillas de ductos de aire

Se obtienen los vectores que permiten conocer, por medio de un análisis estadístico, el comportamiento del aire en las proximidades de las rejillas de los ductos de aire. En la siguiente Figura se presenta un resultado típico de mediciones efectuadas por medio de esta técnica. En ella pueden apreciarse los vectores velocidad de una gran diversidad de puntos sobre el campo de flujo, cada uno de los cuales está asociado a una partícula trazadora

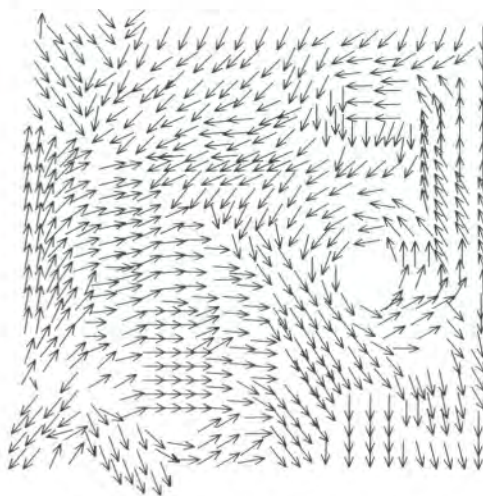


Fig. 7: Mapa de velocidades obtenido a partir de la técnica PIV (caso hipotético).

Referencias

- [1] Buchhave, P. "Optical processing of PIV images". *Journal of Flow Visualization and Image Processing*, vol. 2, pp. 149-160, 1995.
- [2] González, R., Woods, R. Digital image processing. Addison-Wesley Pub. Co. London, 1994.
- [3] Fujita, I., Kaizu, T. "Correction method of erroneous vectors in PIV". *Journal of Flow Visualization and Image Processing*, vol. 2, pp. 173-185, 1995.
- [4] Gray, C. PIV Seminar. University of Edinburgh, Universidad Carlos III and Optical Flow System Ltd. Madrid, 1996.
- [5] Hirt, F., Jud, E., Zhang, Z. "Investigation of the local flow topology in the vicinity of a prosthetic heart valve using PIV". *Proceedings of the Seventh Symposium on Applications of Laser Techniques to Fluid Mechanics*. Lisbon, 1994.
- [6] Holman, J. Métodos experimentales para ingenieros. McGraw-Hill. México D.F., 1986.
- [7] Kimura, I., Kuroe, Y., Ozawa, M. "Application of neural networks to quantitative flow visualization". *Journal of Flow Visualization and Image Processing*, vol. 2, pp. 261-269, 1995.
- [8] Longmire, E., Alashyari, A. "PIV measurements in simulated microbursts". *Proceedings of the Seventh Symposium on Applications of Laser Techniques to Fluid Mechanics*. Lisbon, 1994.
- [9] Okamoto, K. "Three-dimensional particle tracking algorithms: velocity vector histogram and spring model". *Proceedings of The International Workshop on PIV Fukui '95*, pp. 21-32, 1995.
- [10] Russ, J. The image processing handbook. CRC Press. Boca Raton, 1994.